

# **From genome to phenome**

roman.chrast@unil.ch

**The term **phenome** describes the whole organism translation of the genome into cell, tissue, and systemic phenotypes or characteristics.**

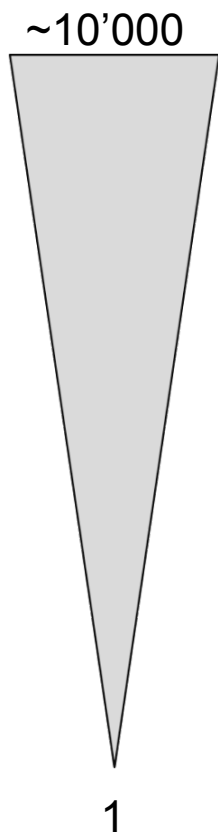
**So in other words:**

**The phenome is what the genome really does.**

# Candidates

# of candidates

function



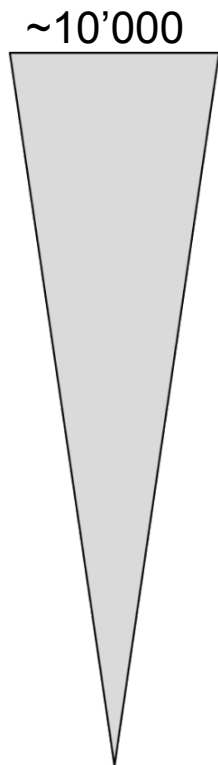
house keeping



specialized

## Candidates

# of candidates



1

function

house keeping

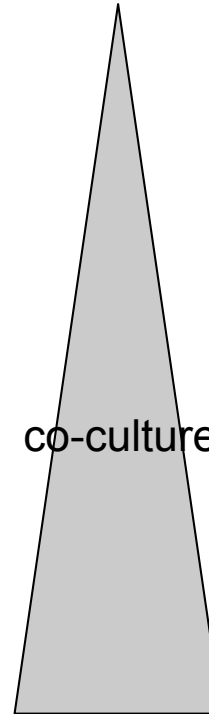


specialized

## Approach

*in vitro*

cells

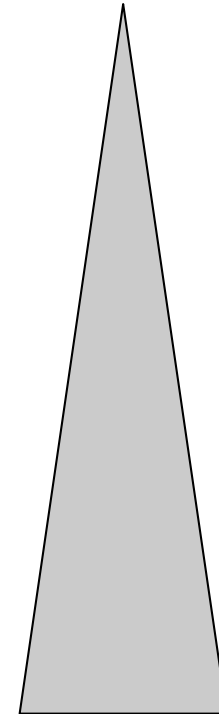


co-culture

explants

*in vivo*

E. Coli



M. Musculus



# Approaches to decipher how the genome specifies the phenome

## **Sequence-driven approaches:**

focusing first on specific DNA sequence variants and then identifying phenotypic consequences

## **Phenotype driven approaches:**

by identifying individuals with a different phenotype and then linking the phenotype to DNA sequence variants

Table 1 | **Attributes of some key animals used to model human disease**

Attribute of disease model	Model organism			
	<i>Fly</i>	<i>Zebrafish</i>	<i>Mouse</i>	<i>Rat</i>
<i>Practical issues</i>				
Husbandry infrastructure	\$	\$	\$\$\$	\$\$\$
Cost per animal per year	\$	\$	\$\$\$	\$\$\$
Characterized inbred strains	+	-	++++	+++
Outbred laboratory strains	+	+++	++	++
Anatomical similarity	-	+	++	++
Molecular or genetic similarity	+	++	+++	+++
Pathological similarity	-	++	+++	+++
Storage; for example, freezing sperm	No	Yes	Yes	Yes
<i>Molecular biology tools</i>				
Transgenesis*	++	++	++	++
Targeted gene modification*	+	-	++++	+
Transient <i>in vivo</i> assays*	++	++++	+	+
Allelic series from TILLING*	+++	++++	++	+
Feasibility of large-scale screens <sup>‡</sup>	++++	+++	++	+
Affordability of large-scale screens <sup>‡</sup>	++++	+++	+	-
Sequencing progress <sup>§</sup>	+++	++	+++	++
Annotation progress <sup>§</sup>	++	++	++++	++
<i>Cell-biology tools</i>				
Cell lines and tissue culture	++	+	++++	+
Antibody reagents	++	+	++++	++

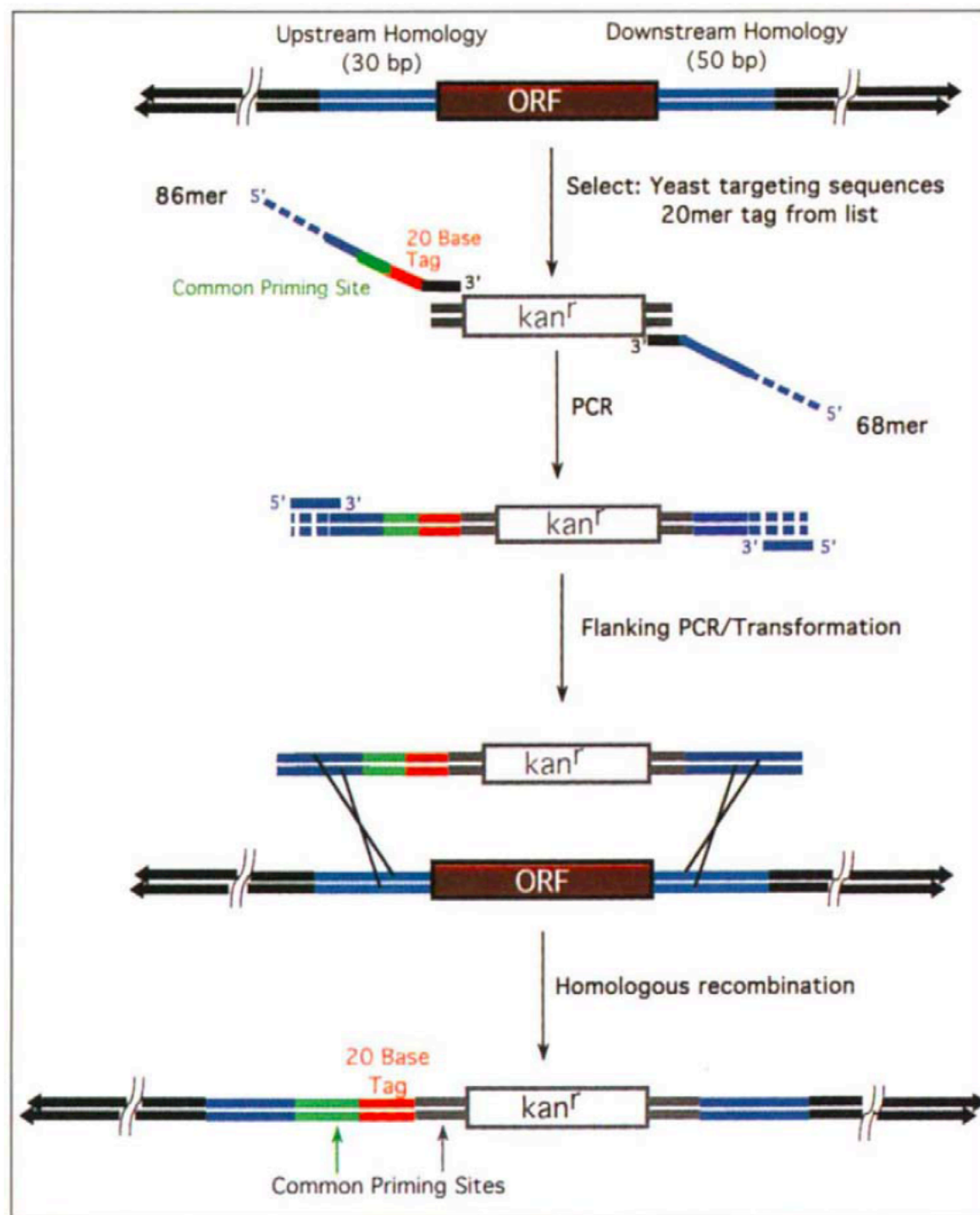
\*Reverse-genetics approach; <sup>‡</sup>forward-genetics approach; <sup>§</sup>genome sequence; -, not relevant, or not a strength; \$, \$\$, \$\$\$ and +, ++, +++, relative cost (\$) and strength (+) of the model in each category; +++++, outstanding strength of the model; TILLING, targeting induced local lesions in genomes.

## **Sequence-driven approaches**



# **Quantitative phenotypic analysis of yeast deletion mutants using a highly parallel molecular bar-coding strategy**

Daniel D. Shoemaker<sup>1</sup>, Deval A. Lashkari<sup>1</sup>, Don Morris<sup>2</sup>, Mike Mittmann<sup>2</sup> & Ronald W. Davis<sup>1</sup>

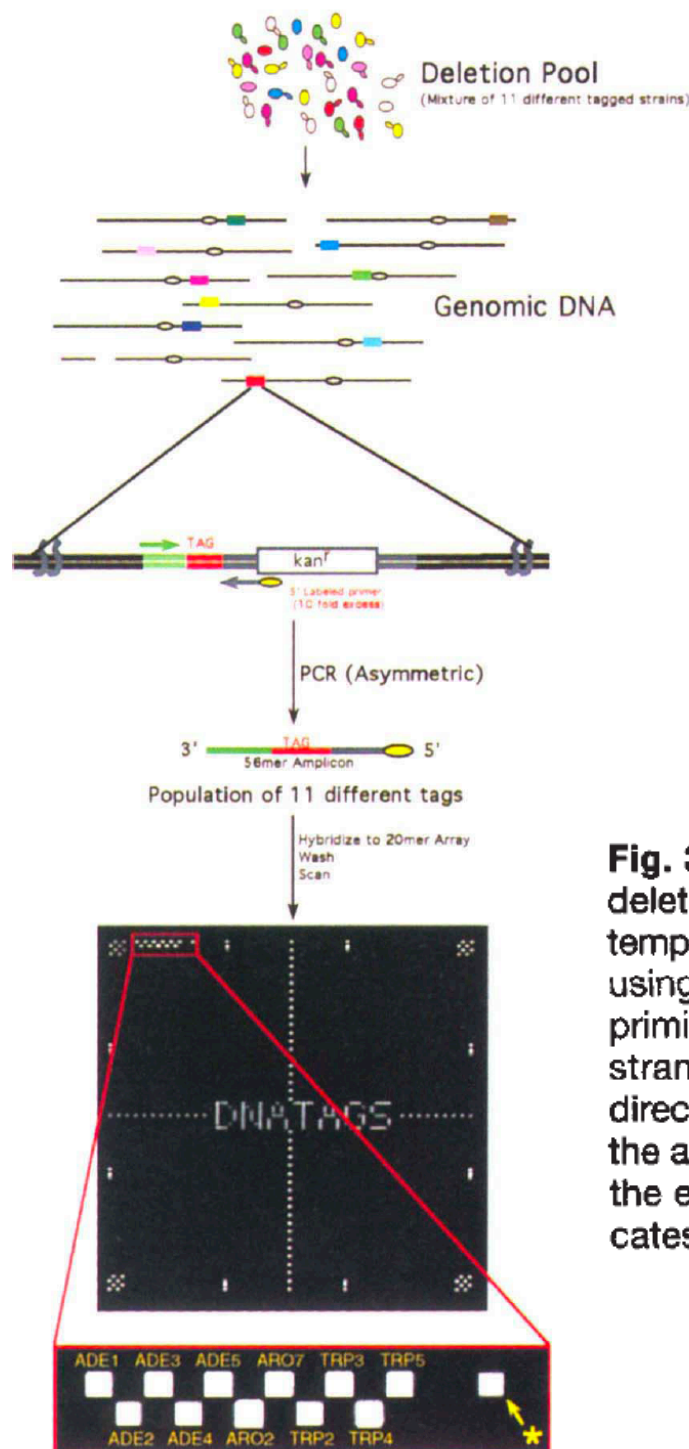


**Fig. 2** PCR-targeting strategy used to generate tagged deletion strains. A dominant selectable marker (*kan<sup>r</sup>*) was amplified using a pair of long primers that contained yeast sequences on the 5' end and homology to the marker on the 3' end. One of the oligonucleotides is a 68mer that contains 50 bases of yeast homology and 18 bases of homology to the marker. The other oligonucleotide is an 86mer that contains a 20 base tag and an 18 base tag priming site in addition to the 30 bases of yeast homology and the 18 bases of marker homology. The dashed lines represent missing sequences on the 5' ends of the long unpurified oligonucleotides (see Methods). A second round of PCR is performed with 20mers that are homologous to the ends of the initial PCR product to increase the amount of full length product. The product from the second round PCR is transformed into a haploid yeast strain and homologous recombination results in the replacement of the targeted ORF with the marker, 20mer tag, and tag priming site.

**Table 1 Transformation results for auxotrophic ORFs**

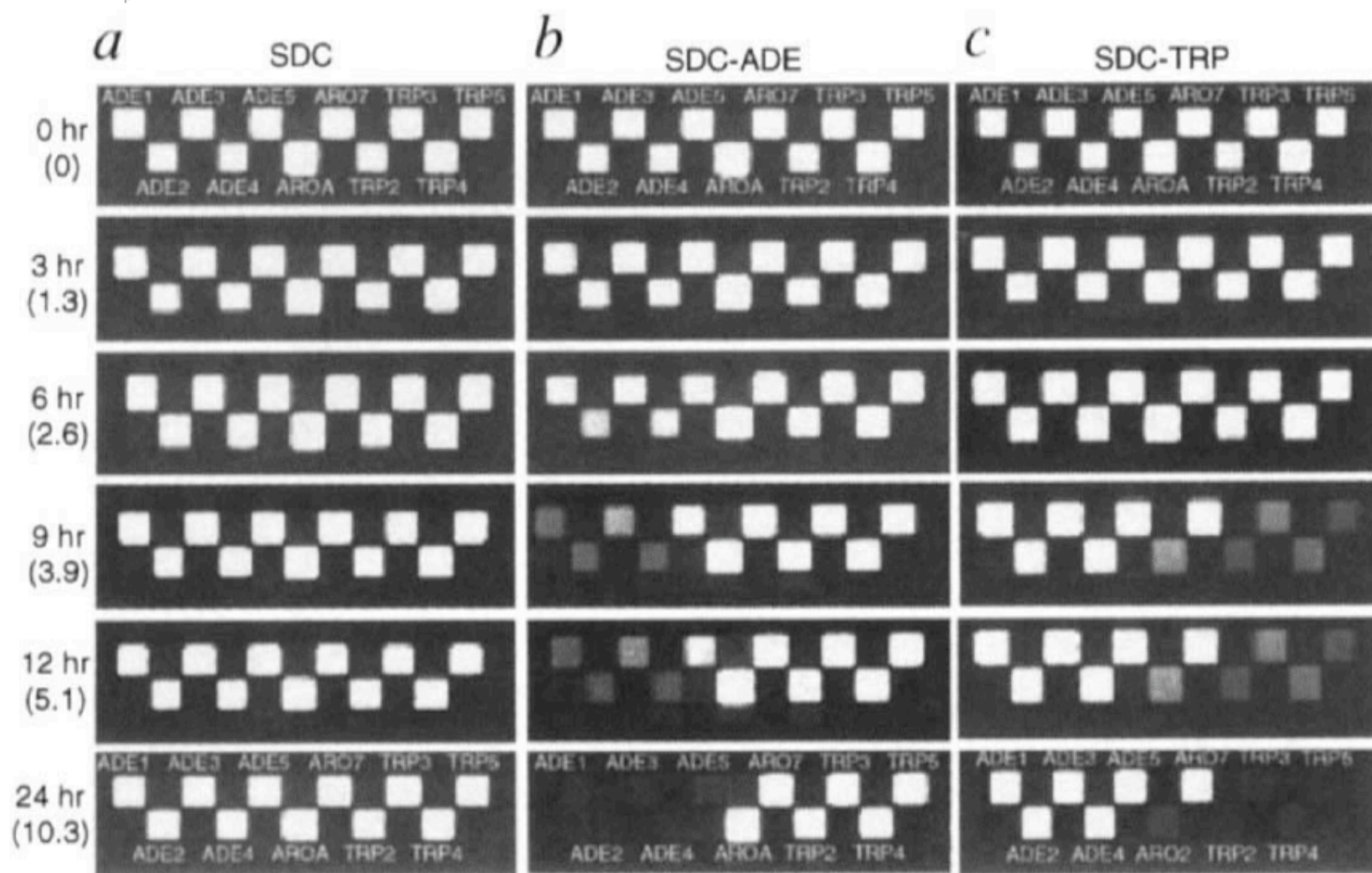
ORF name	ORF size	Targeting efficiency
<i>ADE1</i> (M61209)	305aa	100% (8/8)
<i>ADE2</i> (M59824)	571aa	50% (4/8)
<i>ADE3</i> (M24737)	945aa	50% (4/8)
<i>ADE4</i> (M74309)	509aa	100% (8/8)
<i>ADE5</i> (X04337)	802aa	100% (8/8)
<i>ARO2</i> (X60190)	376aa	75% (6/8)
<i>ARO7</i> (M24517)	256aa	100% (8/8)
<i>TRP2</i> (K01388)	507aa	100% (8/8)
<i>TRP3</i> (K01386)	484aa	88% (7/8)
<i>TRP4</i> (X04273)	380aa	75% (8/8)
<i>TRP5</i> (V01342)	707aa	88% (7/8)

The ORF name (GenBank accession number), ORF size and targeting efficiency (correct integrations versus total) are shown for each of the 11 auxotrophic ORFs.



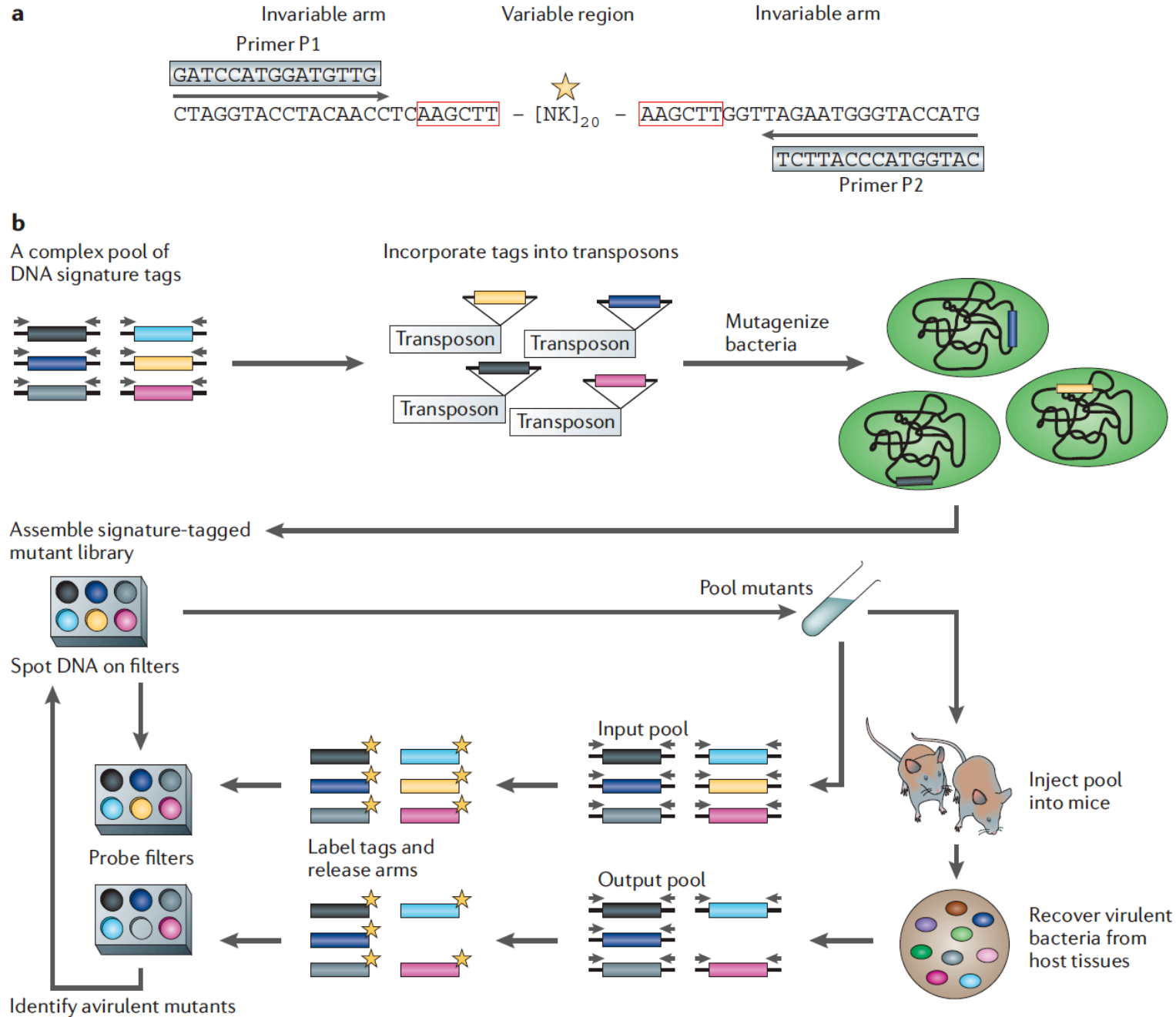
**Fig. 3** Tag amplification strategy. Genomic DNA isolated from a deletion pool containing 11 tagged deletion strains was used as template for the tag amplification reaction. The tags were amplified using a pair of primers that were complementary to the common priming sites. The asymmetric PCR generates a mixture of single-stranded fluorescently labelled 56mer tag amplicons that were directly hybridized to the high-density array. A scanned image of the array is shown. A close-up view of the left hand corner shows the exact location of the 11 tags on the array. The asterisk indicates a cross-hybridizing sequence (see text).



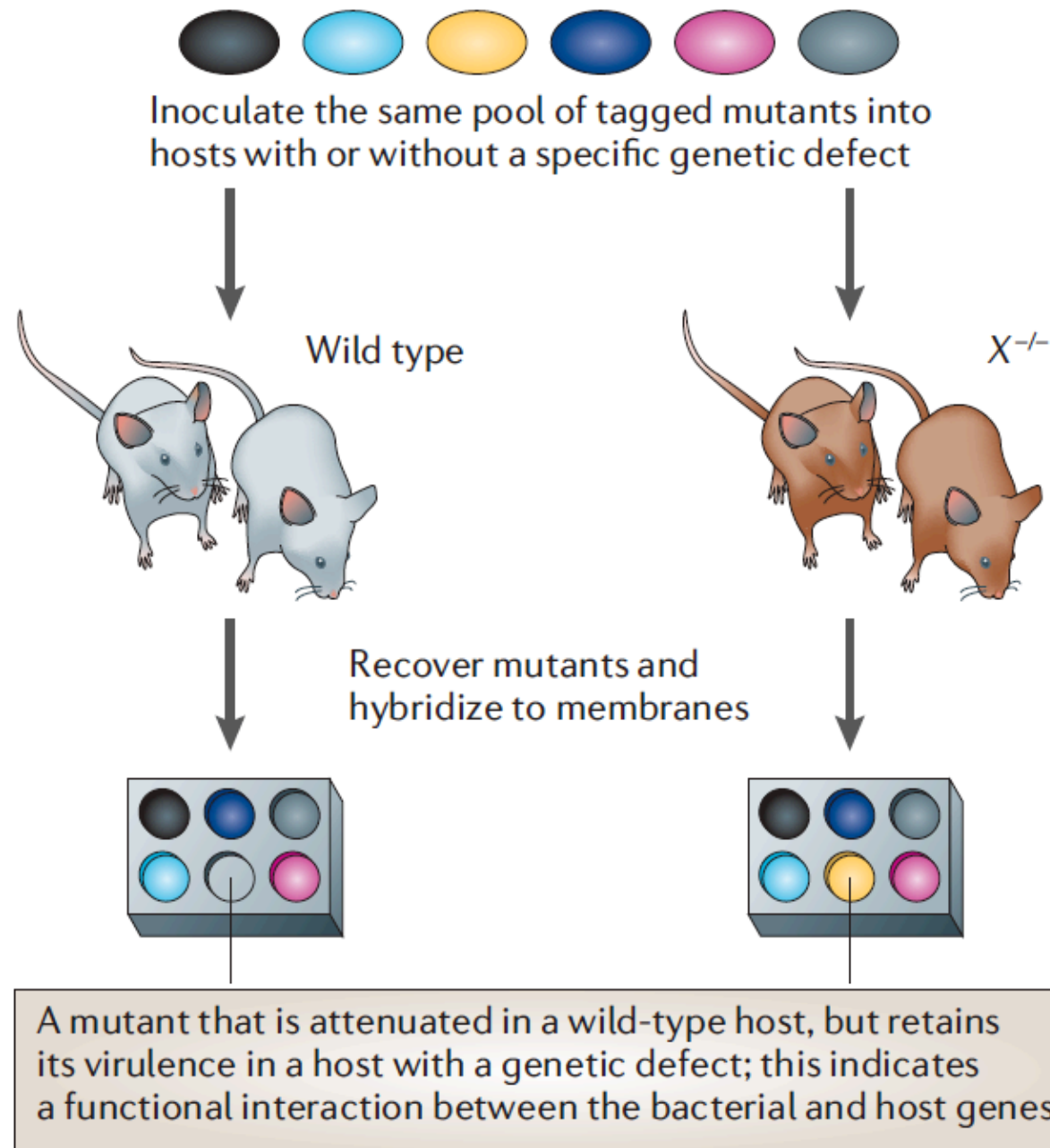




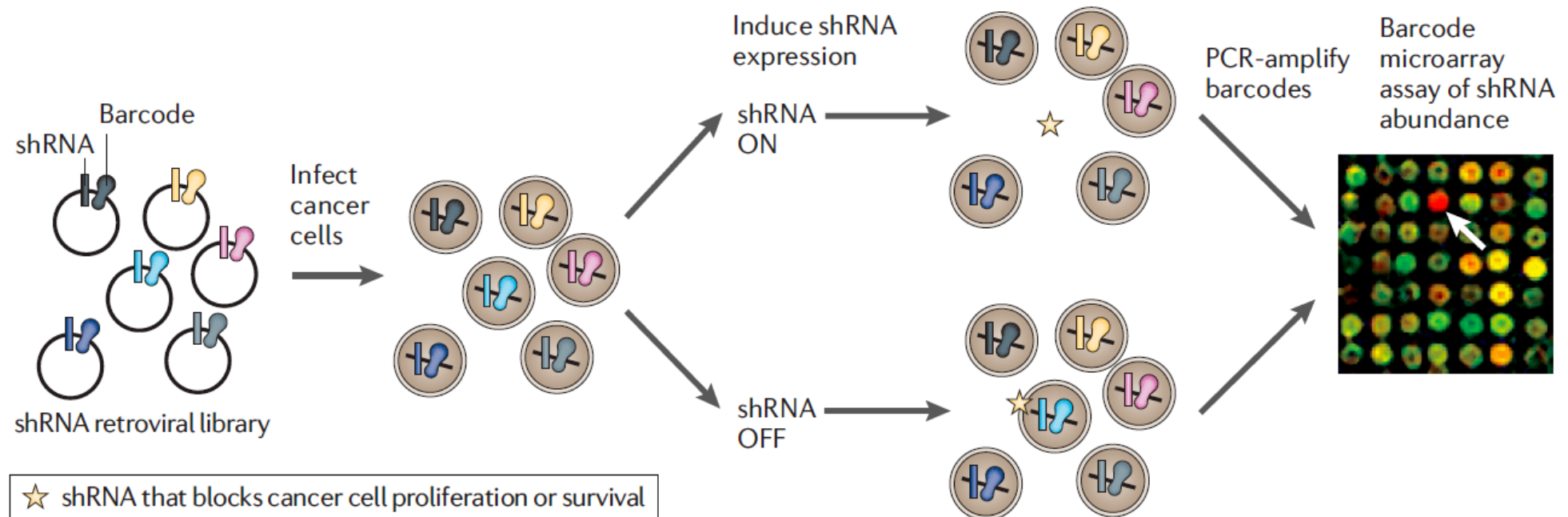
# Signature-tagged mutagenesis of *Salmonella*



# Differential signature-tagged mutagenesis screen

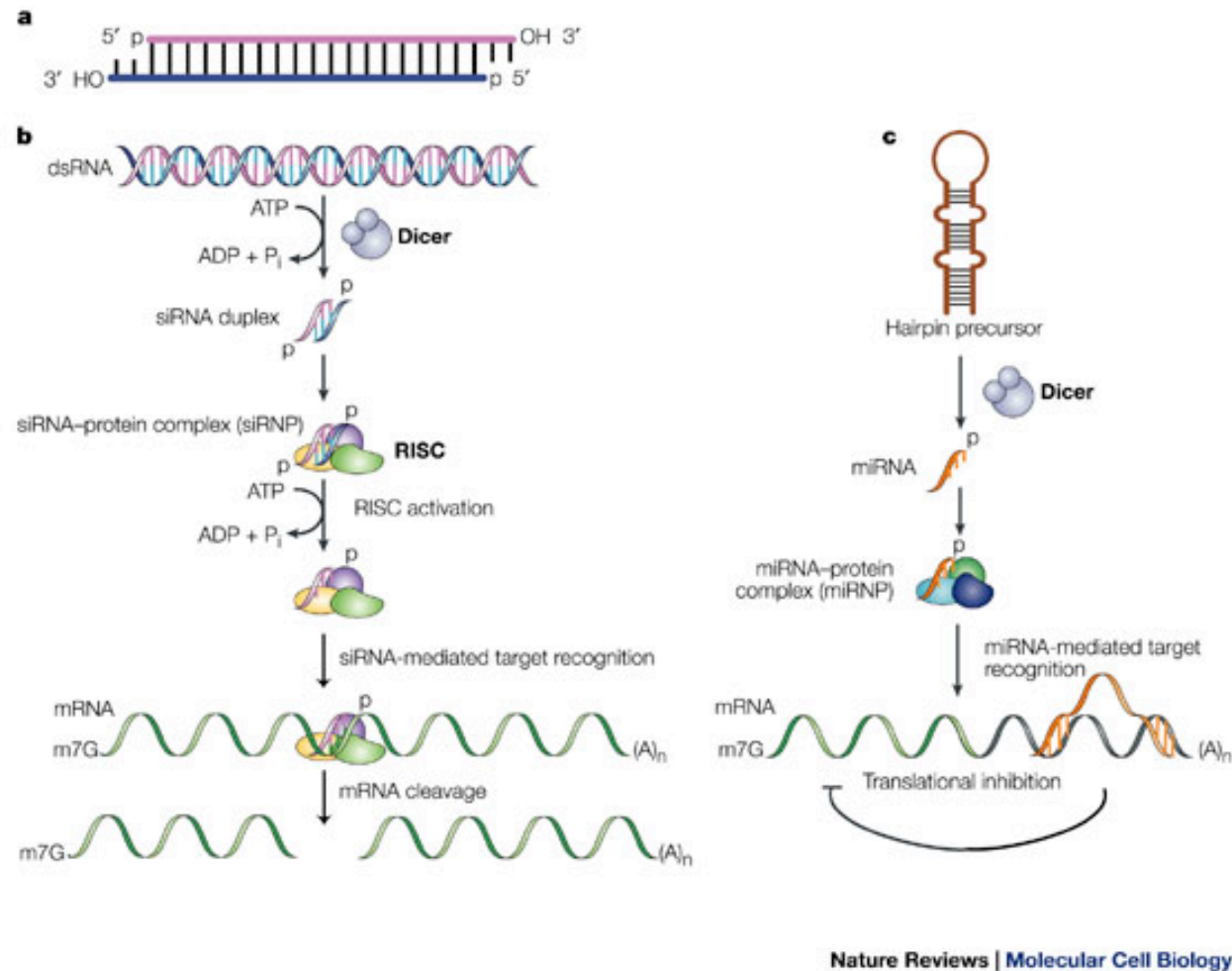


# RNAi screen with barcoding



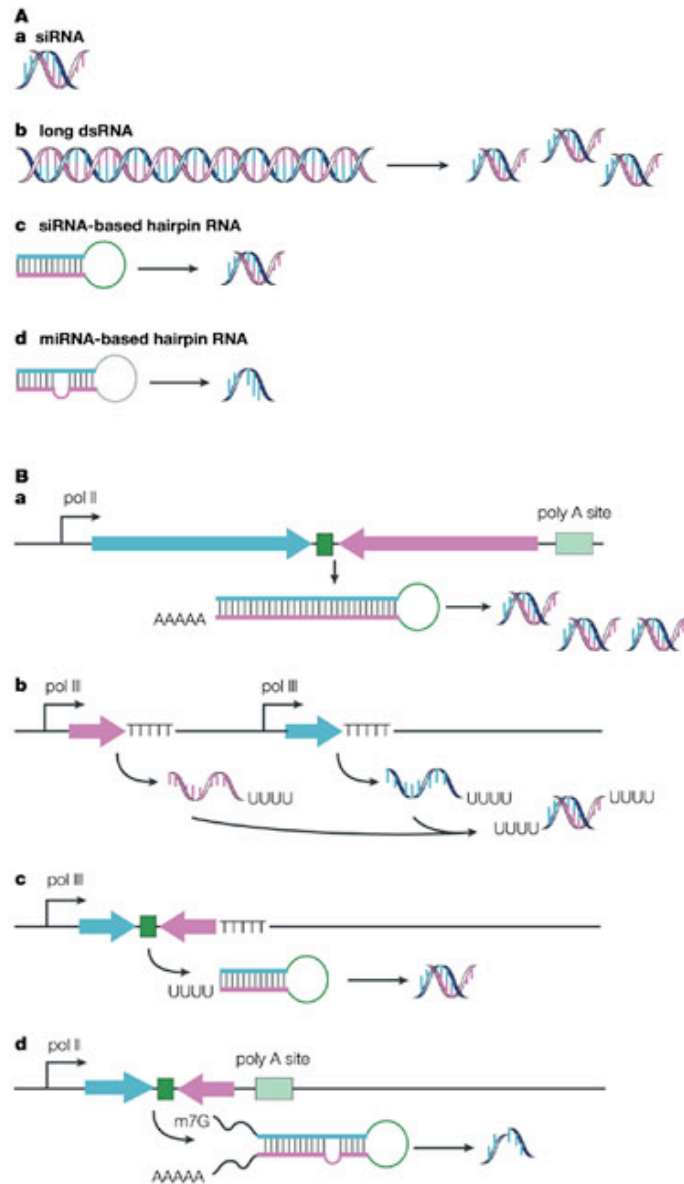
**Figure 5 | RNAi screen with barcoding.** Retroviral vectors that encode a library of short hairpin RNAs (shRNAs) that are under the control of an inducible promoter are introduced into a cancer cell line. The cells are then divided in two subpopulations: one is subjected to induction of shRNA expression and the other is used as a control cell population. An shRNA that reduces the expression of a protein that is critical for proliferation or survival of the cancer cells will be eliminated from the induced, shRNA-expressing culture. Genomic DNA is isolated from the two populations at different time points, and PCR is used to amplify the barcodes that are present in the genomic DNA. Amplified DNAs from the induced and control cultures are labelled with different fluorescent dyes and cohybridized to barcode oligonucleotides on microarrays to determine the relative abundance of each barcode in the two populations. This indicates the relative depletion or enrichment of cells that express a given shRNA<sup>65</sup>.

# The RNA interference pathway



a) Short interfering (si)RNAs. Molecular hallmarks of an siRNA include 5' phosphorylated ends, a 19-nucleotide (nt) duplexed region and 2-nt unpaired and unphosphorylated 3' ends that are characteristic of RNase III cleavage products<sup>14</sup>. **b | The siRNA pathway. c | The micro (mi)RNA pathway.**

# Methods to generate short RNAs that silence gene expression.



## A | Silencing by short RNAs that are generated *in vitro*.

**Aa** | Chemically synthesized short interfering (si)RNAs that are introduced into cells bypass the 'dicing' step and are incorporated into the RNA-inducing silencing complex (RISC) for targeted messenger RNA degradation.

**Ab** | Long double-stranded (ds)RNAs that are introduced into cells can be processed by Dicer into siRNAs that silence gene expression.

**Ac** | Perfect duplex hairpin RNA can be cleaved by Dicer into siRNAs.

**Ad** | Imperfect duplex hairpin RNA, based on pre-micro (mi)RNA structures, can be cleaved by Dicer into miRNAs and direct gene silencing.

## B | Silencing by short RNAs that are generated *in vivo*.

**Ba** | Long hairpin RNA expressed from an RNA polymerase (pol) II promoter yields a population of siRNAs with several sequence specificities. siRNAs with a single sequence specificity can be expressed either by **Bb** | tandem pol III promoters that express individual sense and antisense strands of the siRNA that associate in *trans* or by **Bc** | a single pol III promoter that expresses short hairpin (sh)RNA with the sense and antisense strands of the siRNA that associate in *cis*. **Bd** | Incorporation of an imperfect duplex hairpin structure that is based on pre-miRNA structures can be expressed from a pol II promoter and processed by Dicer into a mature miRNA, which can direct gene silencing.

# LETTERS

---

## **RNAi-mediated gene silencing in non-human primates**

Tracy S. Zimmermann<sup>1</sup>, Amy C. H. Lee<sup>2</sup>, Akin Akinc<sup>1</sup>, Birgit Bramlage<sup>3</sup>, David Bumcrot<sup>1</sup>, Matthew N. Fedoruk<sup>2</sup>, Jens Harborth<sup>1</sup>, James A. Heyes<sup>2</sup>, Lloyd B. Jeffs<sup>2</sup>, Matthias John<sup>3</sup>, Adam D. Judge<sup>2</sup>, Kieu Lam<sup>2</sup>, Kevin McClintock<sup>2</sup>, Lubomir V. Nechev<sup>1</sup>, Lorne R. Palmer<sup>2</sup>, Timothy Racie<sup>1</sup>, Ingo Röhl<sup>3</sup>, Stephan Seiffert<sup>3</sup>, Sumi Shanmugam<sup>1</sup>, Vandana Sood<sup>2</sup>, Jürgen Soutschek<sup>3</sup>, Ivanka Toudjarska<sup>1</sup>, Amanda J. Wheat<sup>2</sup>, Ed Yaworski<sup>2</sup>, William Zedalis<sup>1</sup>, Victor Koteliansky<sup>1</sup>, Muthiah Manoharan<sup>1</sup>, Hans-Peter Vornlocher<sup>3</sup> & Ian MacLachlan<sup>2</sup>



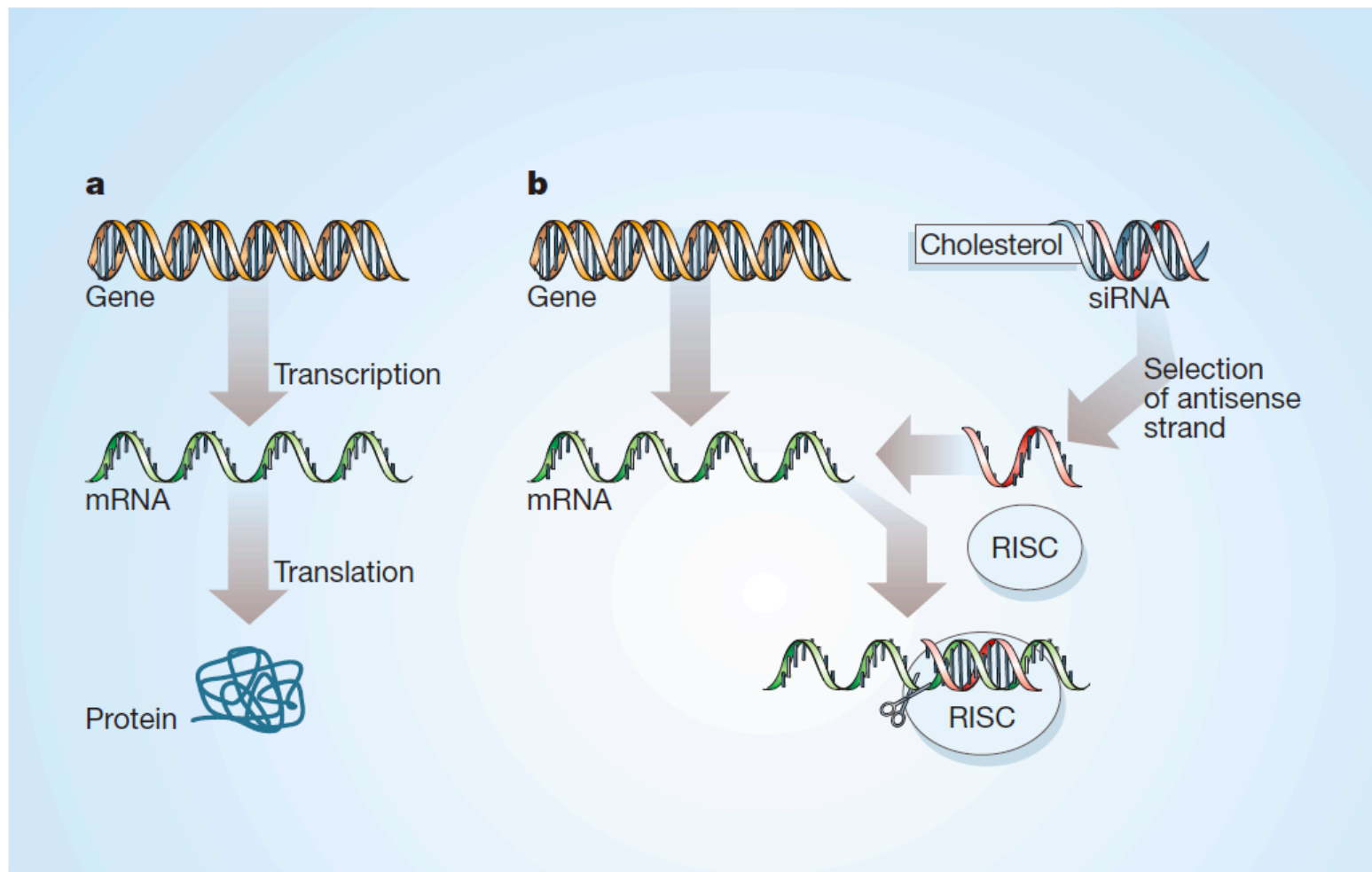
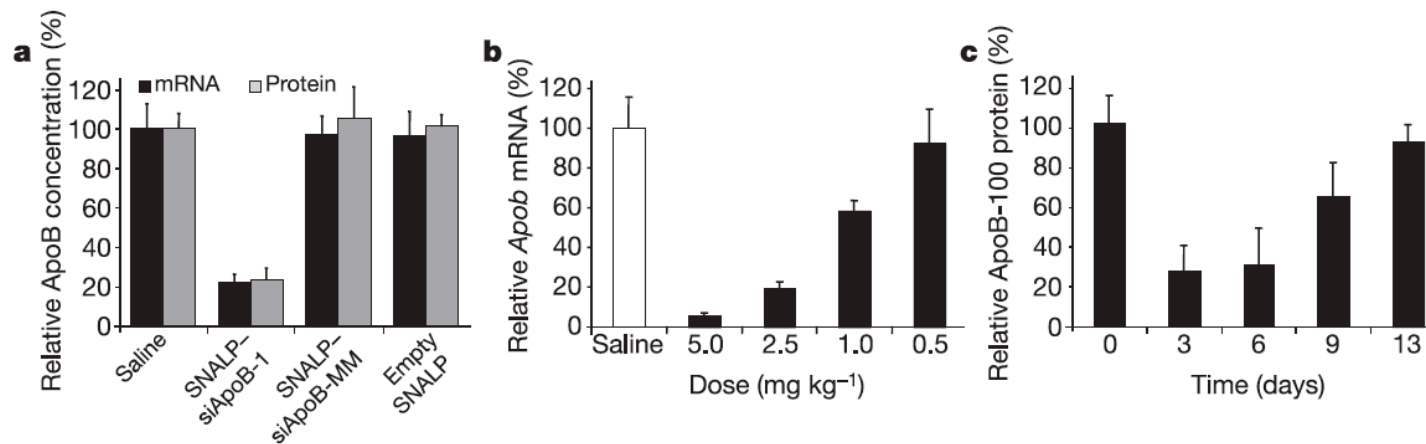


Figure 1 Silencing genes the RNAi way. a, For a gene to be expressed, its DNA sequence must be copied (transcribed) into messenger RNA (mRNA); this must in turn be translated into a protein sequence. b, RNAi works by either destroying the mRNA (bottom) or preventing it from being translated (not shown). In Soutschek and colleagues' modification<sup>1</sup> of the general RNAi approach, short interfering RNAs (siRNAs) are synthesized, chemically modified and labelled on the 'sense' strand (blue) with cholesterol. The siRNAs are then injected intravenously into mice, where the cholesterol group enables the siRNAs to be taken up into tissues. There, the sense strand is destroyed by the inherent RNAi pathway, leaving the antisense strand (red) to bind to a complementary sequence in a target mRNA. Recruitment of a protein complex, the RNA-induced silencing complex (RISC), enables the mRNA to be cleaved.



**Figure 1 | SNALP-siRNA-mediated silencing of murine *Apob* is potent, specific, dose-dependent and long-lasting.** **a**, Liver *Apob* mRNA levels normalized to *Gapdh* mRNA and serum ApoB-100 protein levels measured two days after single i.v. injections of saline, SNALP-siApoB-1 (1 mg kg<sup>-1</sup>), mismatched SNALP-siApoB-MM (1 mg kg<sup>-1</sup>) or empty SNALP vesicles (25 mg kg<sup>-1</sup>) ( $n = 5$  per group). **b**, Liver *Apob* mRNA levels normalized to

*Gapdh* mRNA, assessed three days after i.v. administration of saline or 5, 2.5, 1 or 0.5 mg kg<sup>-1</sup> SNALP-siApoB-2 ( $n = 4$  per group). **c**, Serum ApoB-100 levels after i.v. administration of either saline or 2.5 mg kg<sup>-1</sup> SNALP-siApoB-2 ( $n = 6$  per group). Serum ApoB-100 levels for SNALP-siApoB-2-treated animals are relative to the saline-treated group for the same time point. Data show mean  $\pm$  s.d.



# Some vital statistics of the European house mouse (*Mus musculus*)

---

## **Genome** (*Nature*, 2002; 420:520-62)

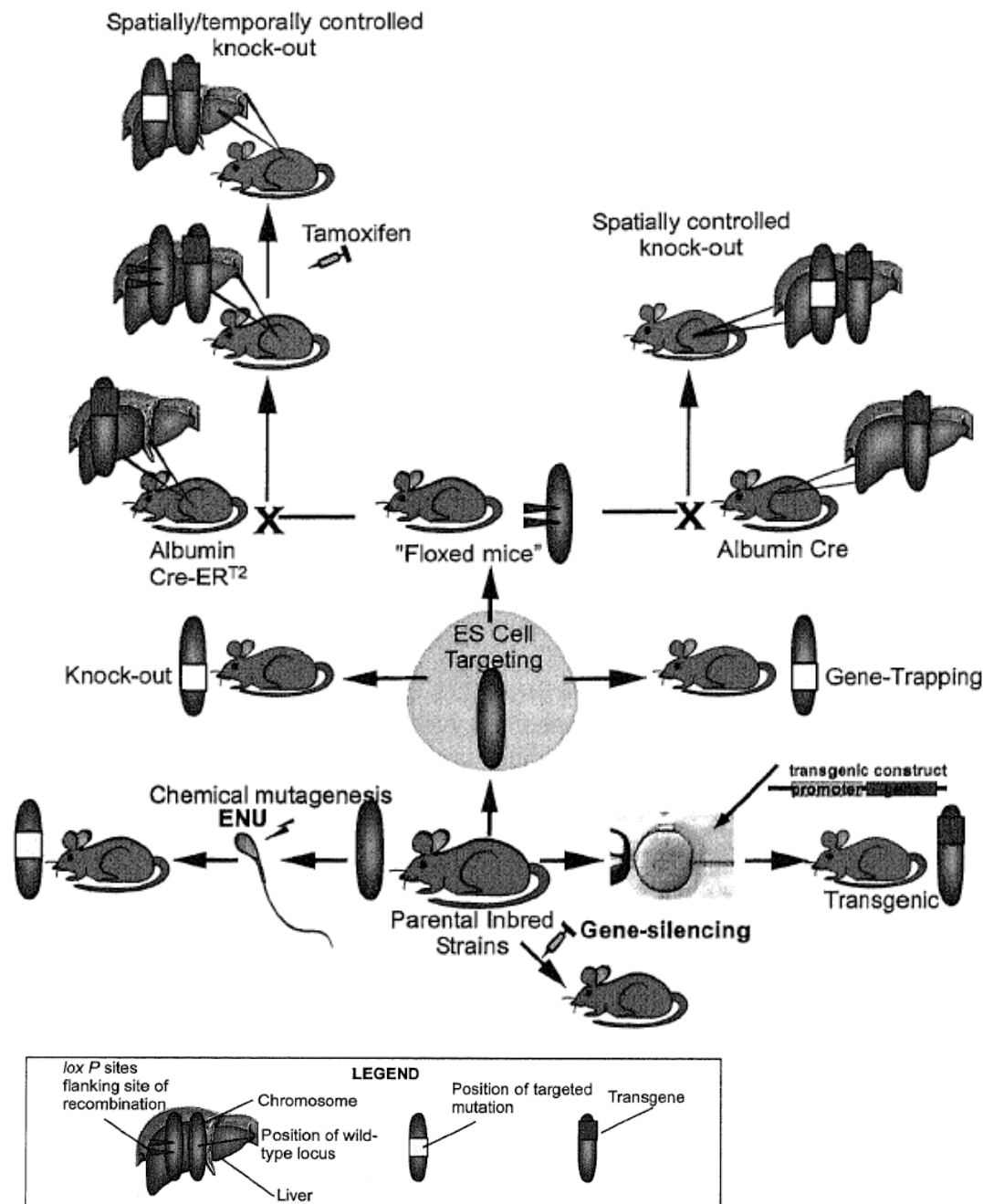
Number of chromosomes	40
Genome size	$2.6 \times 10^9$ bp
Number of genes	22-28'000

## **Reproductive biology**

Gestation time	19-20 days
Age at weaning	3 weeks
Age of sexual maturity	$\geq 6$ weeks
Approx. weight	1g @ birth, 8-12 g @ weaning
Average litter size	6-8
Numbers of litters / breeding female	4-8

Life span in laboratory	1.5 - 2.5 years
-------------------------	-----------------

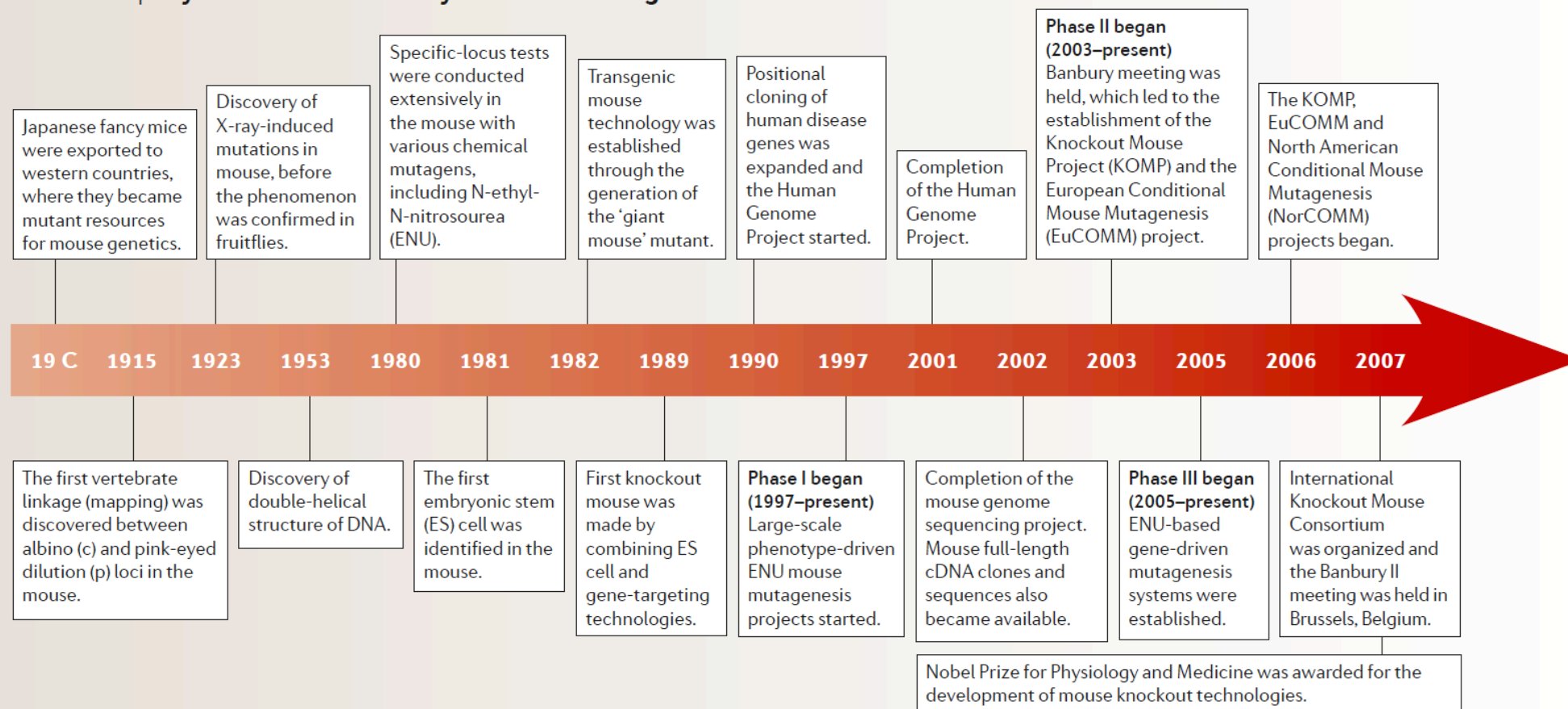
# Genetically engineered mouse models



**Figure 1.** Genetically engineered mouse models (GEMMs)

Gene targeting can be restricted in a tissue and/or temporal manner through generation of "premutant" mice (e.g., "floxed" or flanked with loxP sites) which are then bred with transgenic mice that express the corresponding DNA recombinase (e.g., Cre), allowing for gene deletion to occur in a given tissue (right top panel) and at a given time (left top panel). Gene-trap mutagenesis involves the random insertion of gene-trap vectors into the genome, whereas ethylnitrosourea (ENU) is used as an in vivo mutagen to generate point mutations. Gene silencing is possible through RNAi, a method which uses siRNAs or shRNAs, which can be delivered either by direct administration or via transgenic, homologous recombination, or viral delivery technologies, respectively.

## Timeline | Key events in the history of mouse mutagenesis





## The Nobel Prize in Physiology or Medicine 2007

"for their discoveries of principles for introducing specific gene modifications in mice by the use of embryonic stem cells"



**Mario R. Capecchi**  
University of Utah  
Salt Lake City, UT  
USA

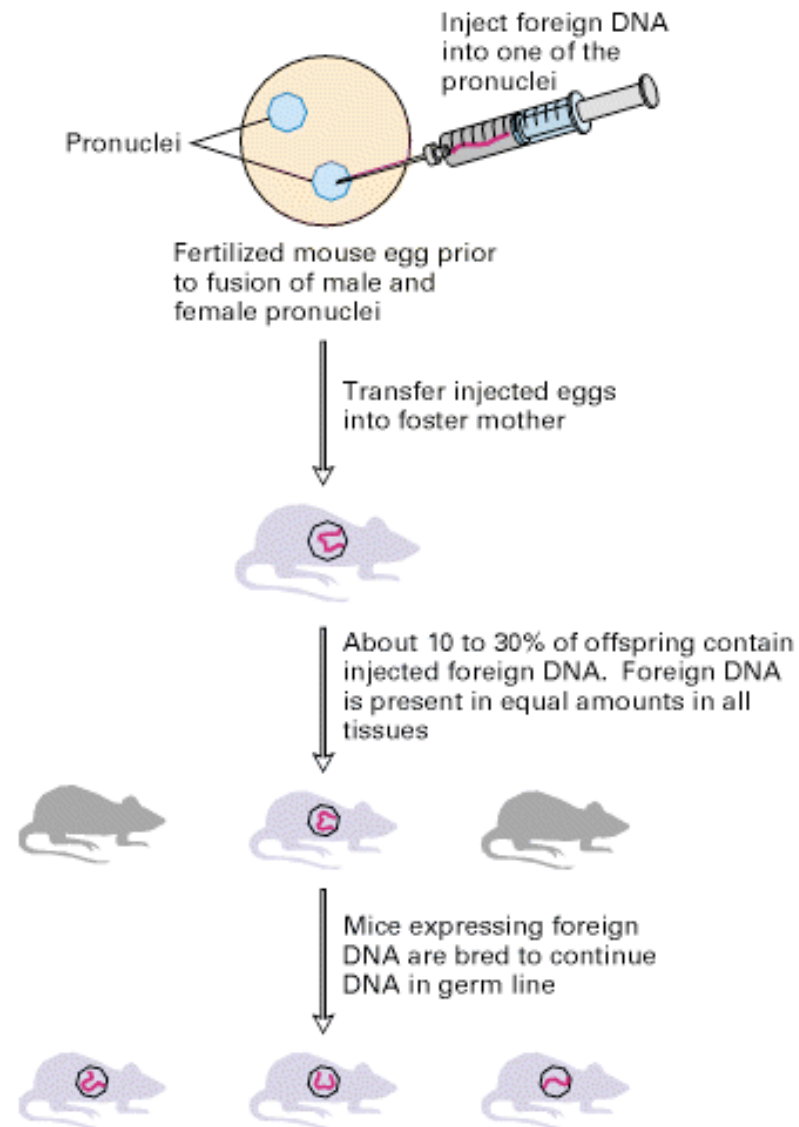


**Sir Martin J. Evans**  
Cardiff University  
Cardiff  
United Kingdom

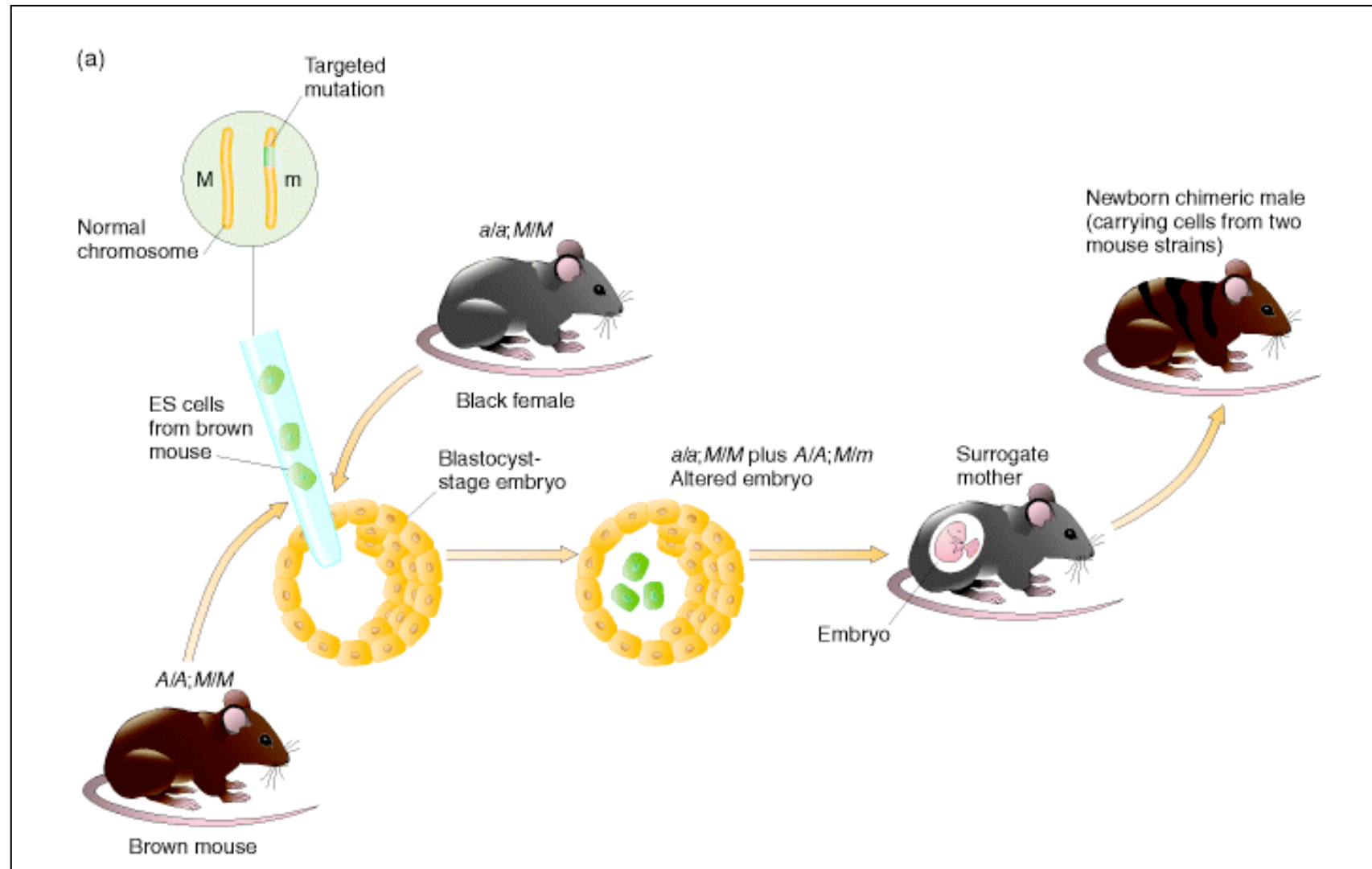


**Oliver Smithies**  
University of North  
Carolina at Chapel  
Hill, Chapel Hill, NC  
USA

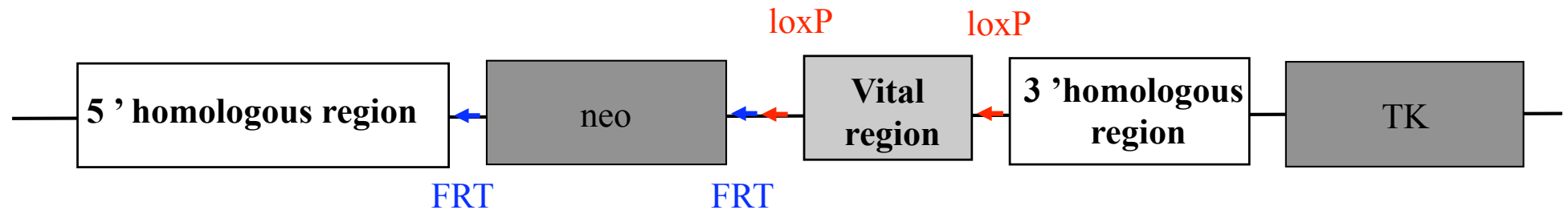
# General procedure for producing transgenic mice



# General procedure for producing KO mice



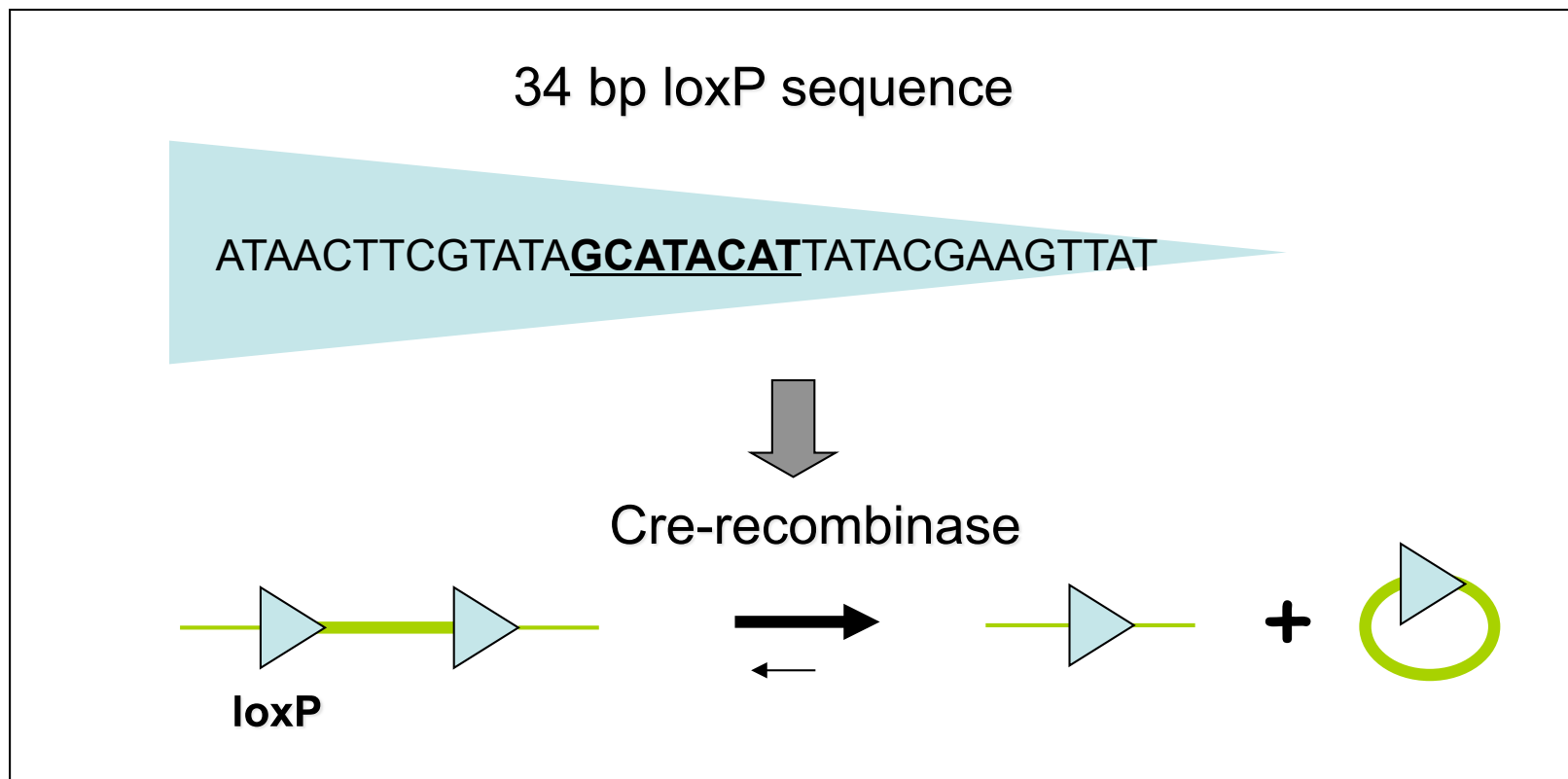
# Construction of a prototypic targeting vector (I)



Targeting vector:

- \* 5' and 3' homologous regions
- \* a vital region flanked by **loxP** sites
- \* a neomycin-resistance (neo) expression cassette (positive selection marker) flanked by two **FRT** sites inserted immediately upstream of the first loxP site
- \* a negative selectable marker (herpes simplex virus thymidine kinase gene) positioned outside the area of genomic homology spliced out when homology recombination occurs

# The Cre/loxP system





# Site-specific recombinases & recognition sequences

Table 1. Alternative SSR Recognition Sites

Target Site <sup>a</sup>	Inverted repeat 1	Spacer	Inverted Repeat 2	Use of Alternative Site	Reference
<i>loxP</i>	ATAACTTCGTATA	ATGTATGC	TATACGAAGTTAT		Hoess, et al., 1982
<i>lox511<sup>b</sup></i>	ATAACTTCGTATA	ATGTATaC	TATACGAAGTTAT	RMCE <sup>c</sup>	Hoess, et al., 1986 Bethke and Sauer, 1997
<i>lox5171</i>	ATAACTTCGTATA	ATGTgTaC	TATACGAAGTTAT	RMCE	Lee and Saito, 1998
<i>lox2272<sup>d</sup></i>	ATAACTTCGTATA	AaGTATcC	TATACGAAGTTAT	RMCE	Lee and Saito, 1998
<i>m2</i>	ATAACTTCGTATA	AgaaAcca	TATACGAAGTTAT	RMCE	Langer, et al., 2002
<i>lox71</i> (LE <sup>e</sup> mutant)	taccgTTCGTATA	ATGTATGC	TATACGAAGTTAT	Stable insertion or inversion	Albert, et al., 1995
<i>lox66</i> (RE <sup>f</sup> mutant)	ATAACTTCGTATA	ATGTATGC	TATACGAaCggta	Stable insertion or inversion	Albert, et al., 1995
<i>FRT</i>	GAAGTTCCTATTC	TCTAGAAA	GTATAGGAACTTC		McLeod, et al., 1986
<i>F<sub>3</sub></i>	GAAGTTCCTATTC	TtcAaAtA	GTATAGGAACTTC	RMCE	Schlake and Bode, 1994
<i>F<sub>5</sub></i>	GAAGTTCCTATTC	TtcAaAAg	GTATAGGAACTTC	RMCE	Schlake and Bode, 1994
<i>FRT</i> mutant -10 (LE mutant)	GAAGTTCaTATTC	TCTAGAAA	GTATAGGAACTTC	Stable insertion or inversion	Senecoff, et al., 1988
<i>FRT</i> mutant +10 (RE mutant)	GAAGTTCCTATTC	TCTAGAAA	GTATAtGAACTTC	Stable insertion or inversion	Senecoff, et al., 1988
<i>attB<sup>g</sup></i>	TCGAGTGAGGTGGAGTACGCGCCCGGGGAGCC CAAGGGCACGCCCTGGCACCCGCA			RMCE or stable inversion	Belteki et al., 2003
<i>attP</i>	CTAGACCCTACGCCCCCAACTGAGAGAACTCAAAGGT TACCCAGTTGGGGCACG			RMCE or stable inversion	Belteki et al., 2003

<sup>a</sup>Nucleotides that have been altered from the wild-type sites are lowercase.

<sup>b</sup>The alternative Cre target site *lox511* (Bethke and Sauer, 1997) is also referred to as *L1* (Hoess et al., 1986) and mutant 71 (Lee and Saito, 1998). *lox511* can recombine, albeit at low efficiency (1–5%), with a wild-type *loxP* site and therefore may have limited use in RMCE reactions (Kolb, 2001, and see text).

<sup>c</sup>RMCE, recombinase mediated cassette exchange.

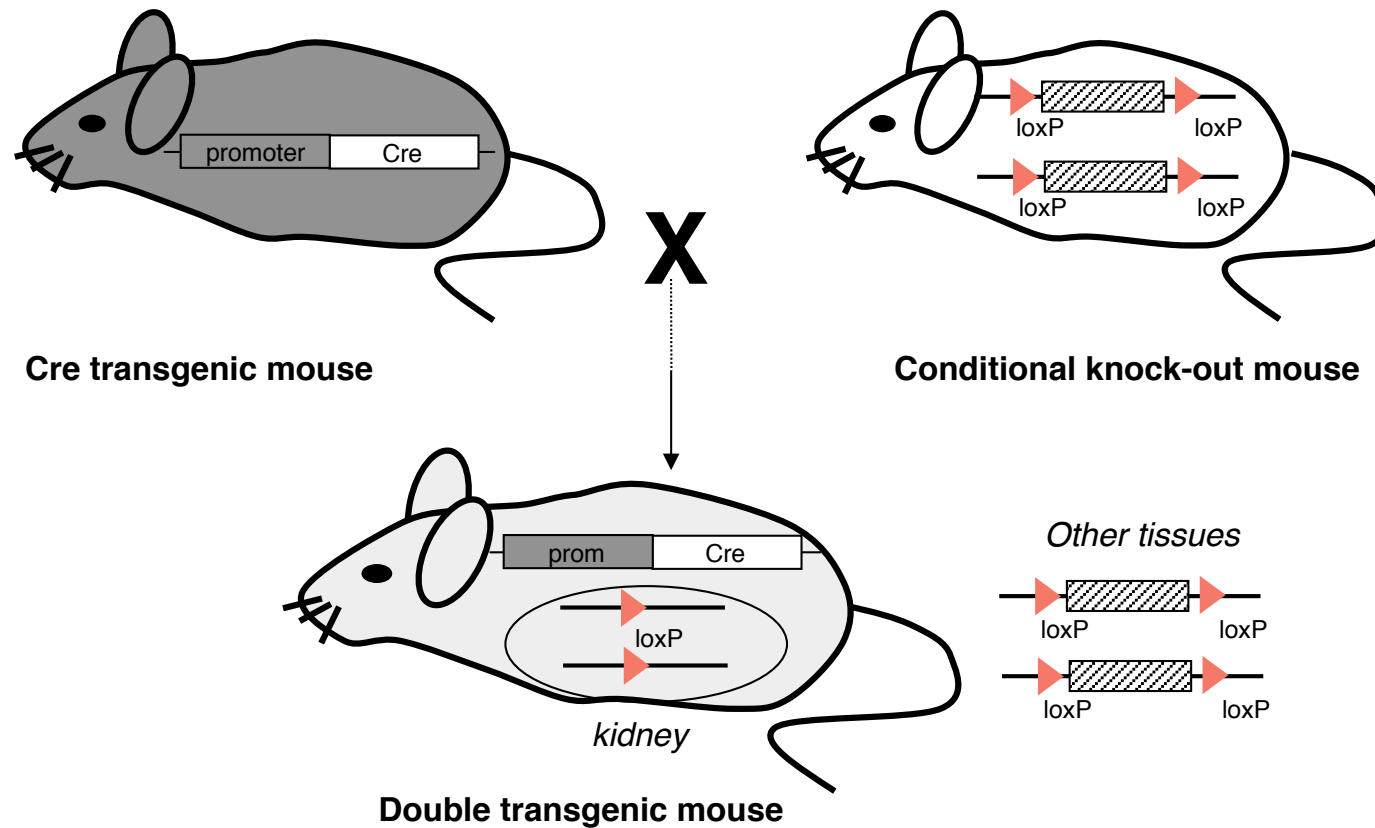
<sup>d</sup>*lox2272* is also referred to as *lox2722* (Kolb, 2001).

<sup>e</sup>LE, left-hand element (LE) (inverted repeat 1) of SSR recognition site.

<sup>f</sup>RE, right-hand element (RE) (inverted repeat 2) of SSR recognition site.

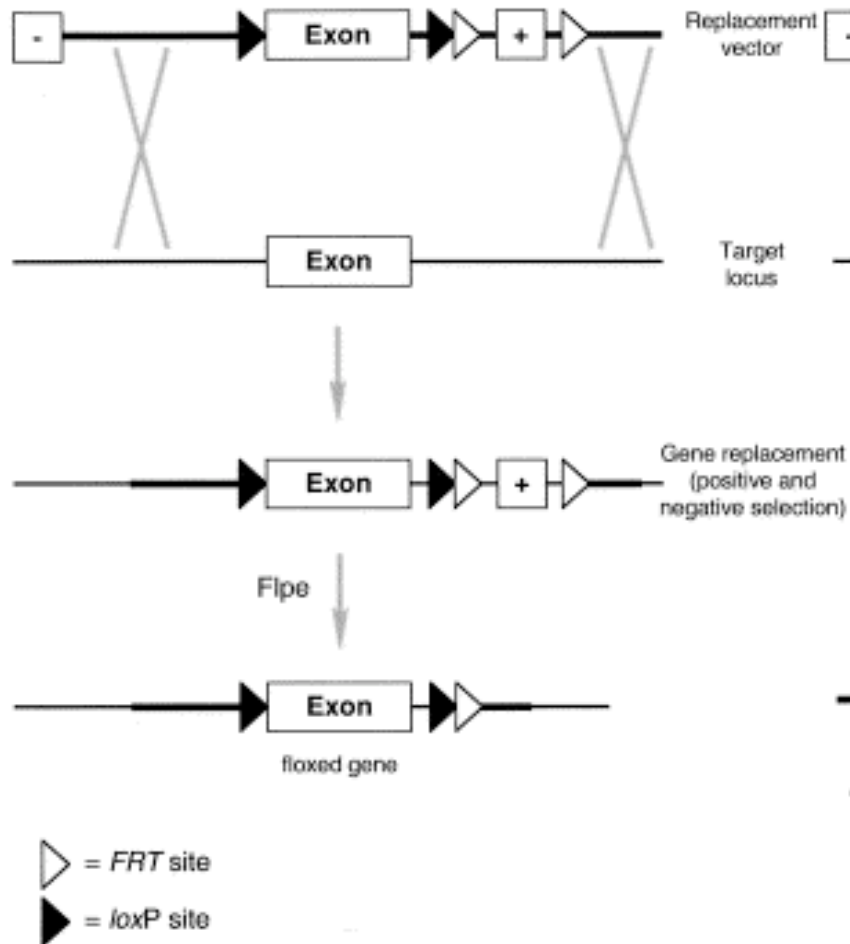
<sup>g</sup>Heterotypic target sites of the  $\phi$ C31 SSR are included here for reference. Core nucleotide triplets (CAA) where the  $\phi$ C31-mediated recombination occurs are underlined.

# Conditional gene targeting: tissue-specificity

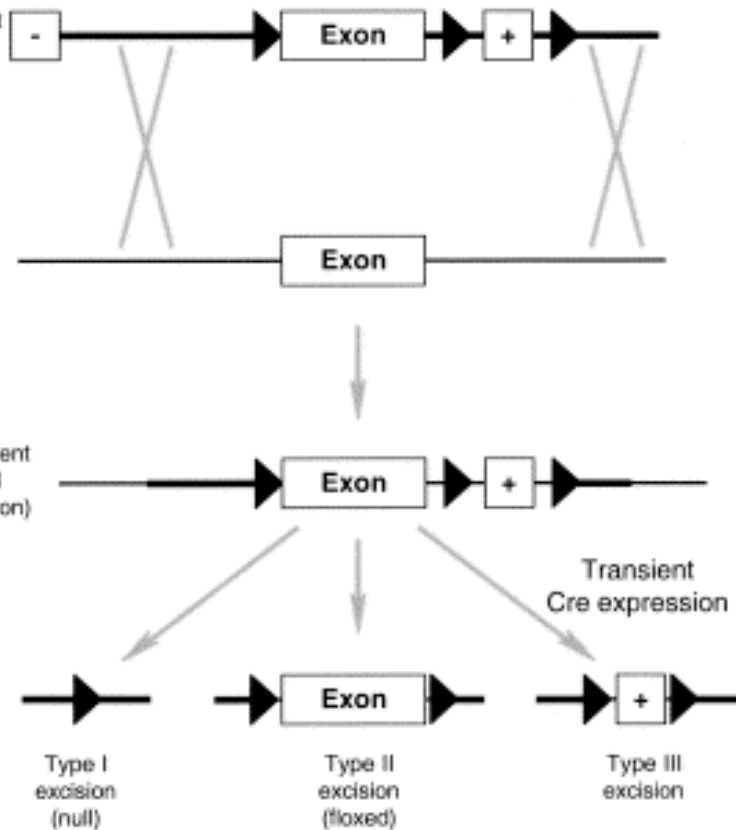


# Construction of a prototypic targeting vector (II)

**A Dual-recombinase strategy**



**B Tri-lox strategy**



## **Available lines (including Cre lines)**

Jackson Laboratory

[http://www.jax.org/resources/mouse\\_resources.html](http://www.jax.org/resources/mouse_resources.html)

Mutant Mouse Regional Resource Centers (MMRRC)

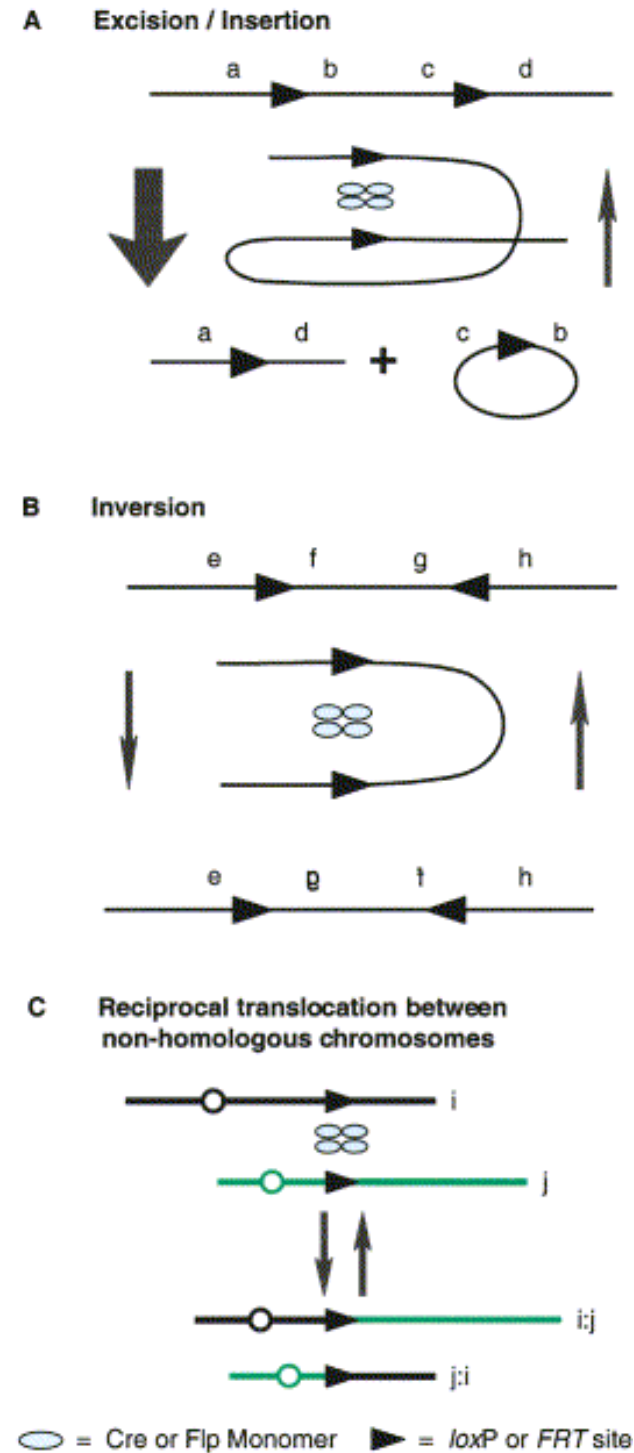
<http://www.mmrrc.org/>

The European Mouse Mutant Archive (EMMA)

[www.emmanet.org/](http://www.emmanet.org/)

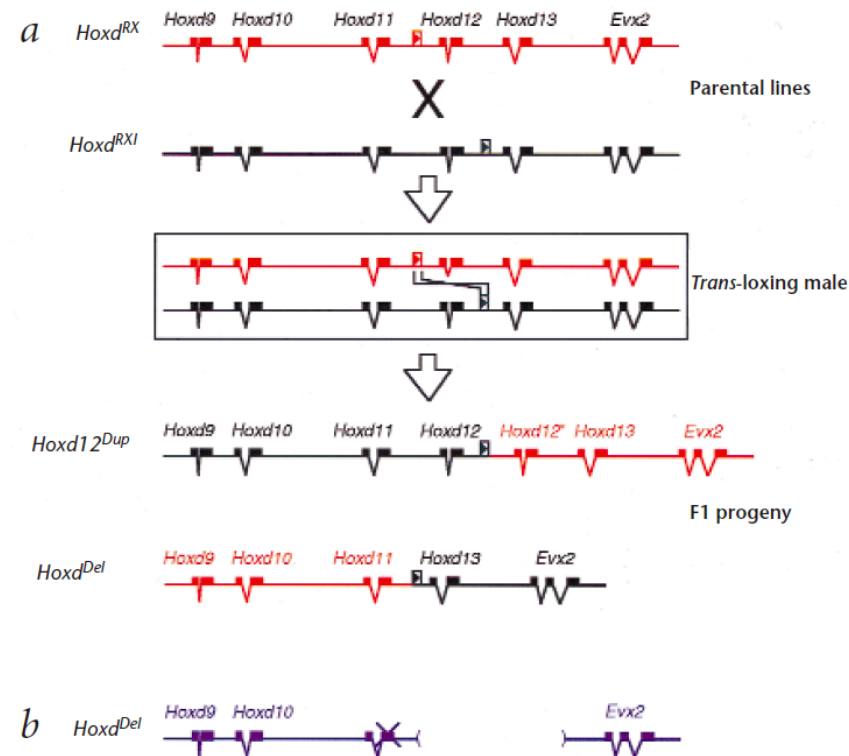
.... ask around

# Complex genomic modifications

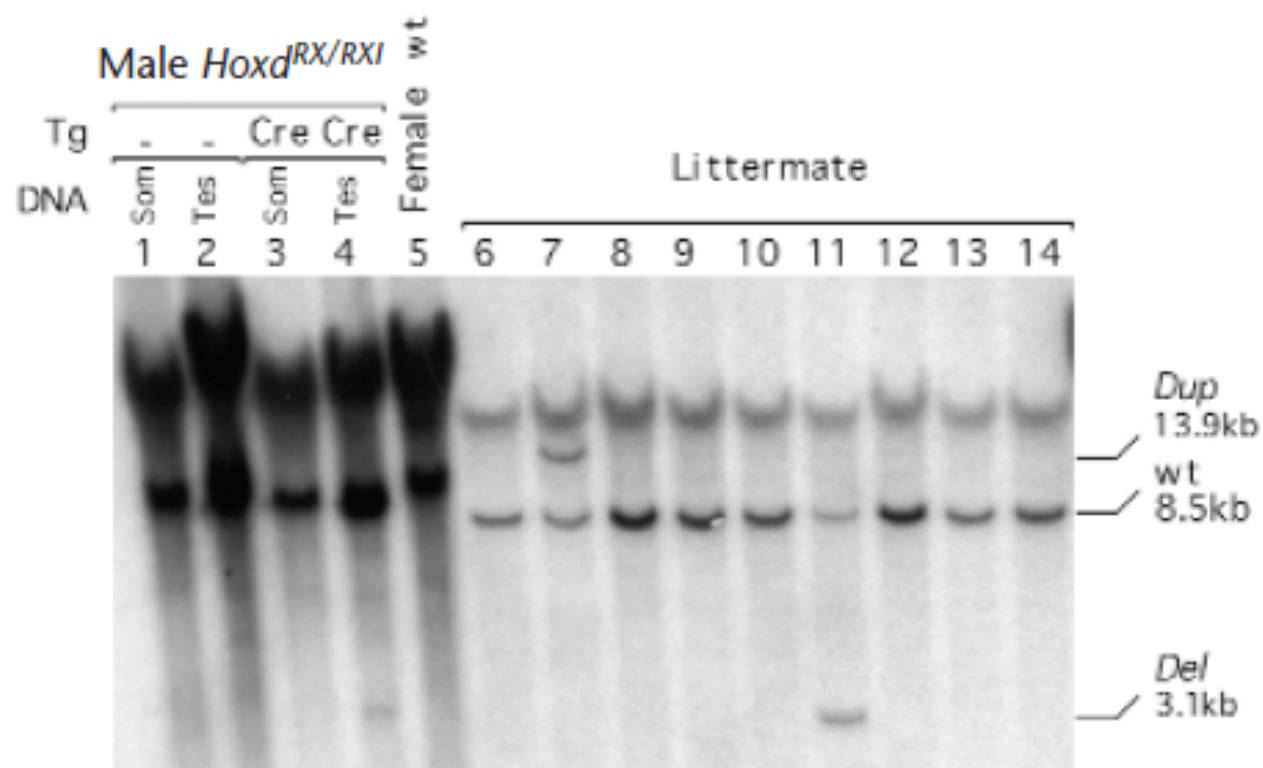


## Engineering chromosomes in mice through targeted meiotic recombination (TAMERE)

Yann Hérault<sup>1</sup>, Minoo Rassoulzadegan<sup>2</sup>, François Cuzin<sup>2</sup> & Denis Duboule<sup>1</sup>



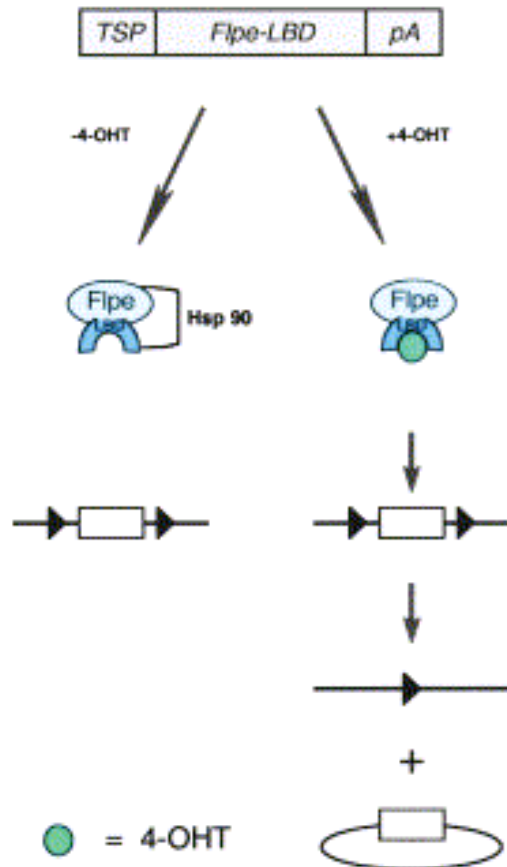
**Fig. 1** The TAMERE strategy. **a**, Schematic representation of Cre-mediated targeted meiotic recombination. The upper two lines indicate the two alleles used for recombination. The *Hoxd<sup>RXI</sup>* locus<sup>2</sup> (black line) has a *loxP* site between *Hoxd12* and *Hoxd13*, whereas the *Hoxd<sup>RX</sup>* locus<sup>16</sup> (red line) carries a *loxP* site between *Hoxd11* and *Hoxd12*. Following chromosome pairing in meiotic prophase and Cre expression through the *Sycp1/Cre* transgene, *trans*-allelic recombination occurs (middle), generating two novel alleles (bottom lines). In the first allele (*Hoxd12<sup>Dup</sup>*), a supernumerary *Hoxd12* is produced (asterisk) in the same orientation as the original. In the second allele (*Hoxd12<sup>Del</sup>*), the *Hoxd12* locus is deleted. **b**, Schematic of the deficiency used to analyse the newly produced alleles. The *Hoxd<sup>Del</sup>* allele is a targeted deletion of *Hoxd12* and *Hoxd13*, whereas *Hoxd11* is inactivated by an in-frame *lacZ* fusion<sup>11</sup>.



# Conditional gene targeting: spatial & temporal specificity (leakiness problems!)

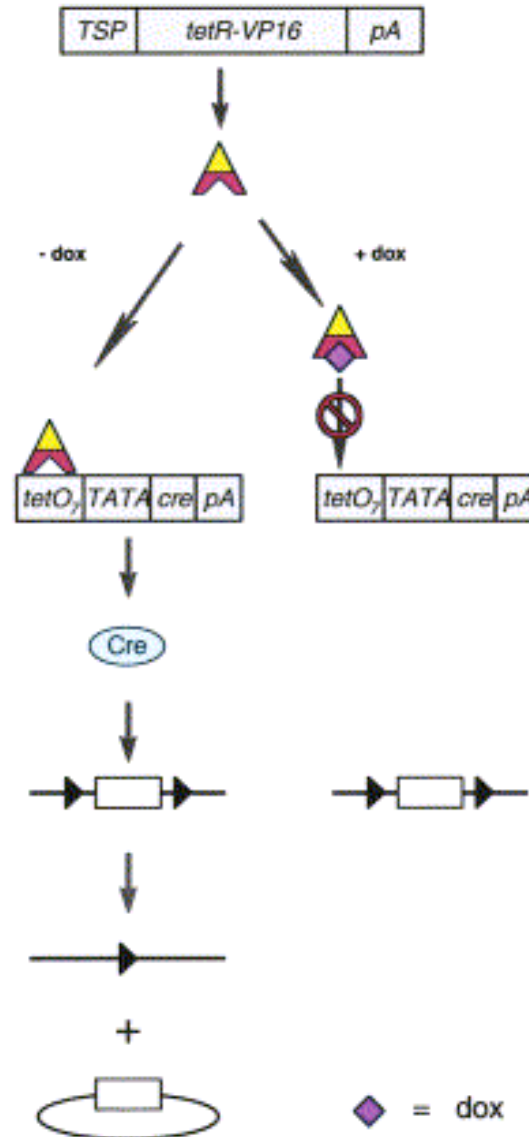
**A Post - translational control**

ligand binding domain (LBD)  
of a mutant estrogen receptor

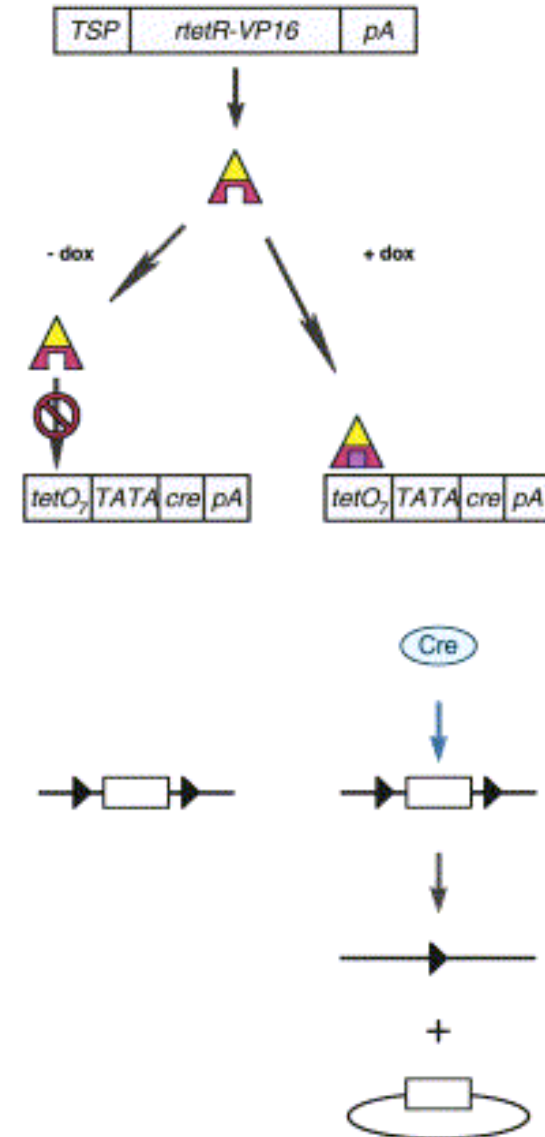


4-hydroxy tamoxifen

**B Transcriptional control - tTA**



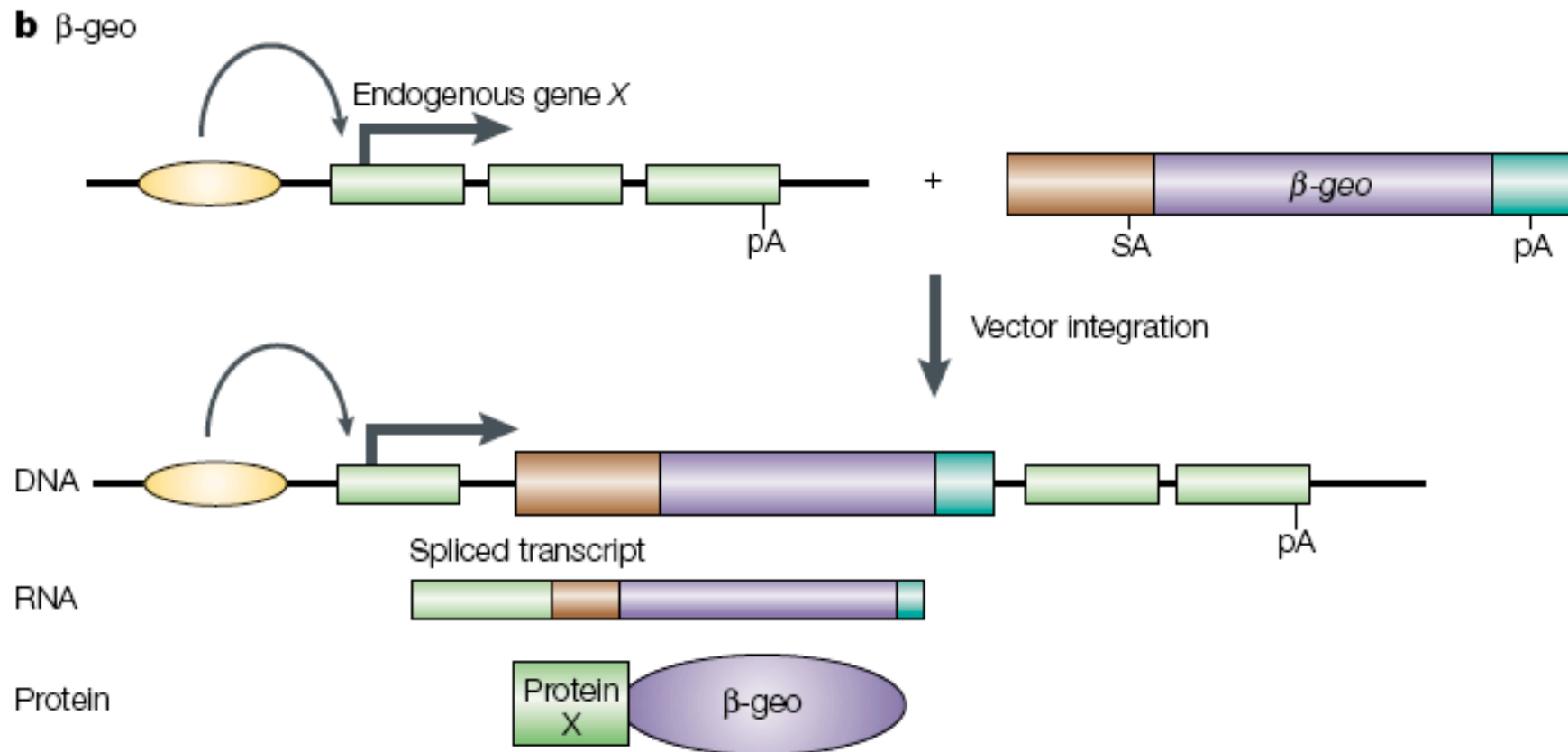
**C Transcriptional control - rtTA**



Branda & Dymecki, Dev. Cell 2004



# Gene trapping in ES cells

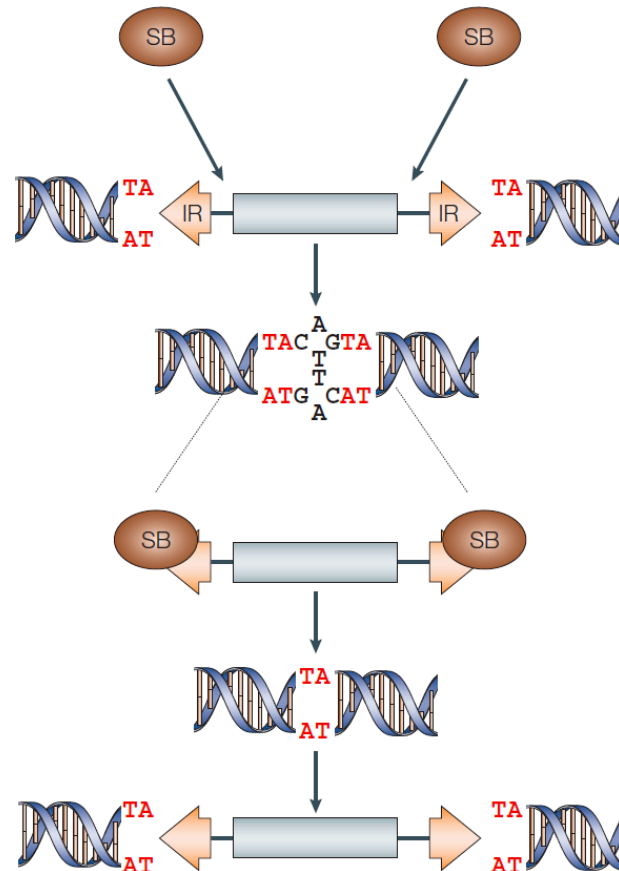


# Sleeping Beauty

## Box 2 | The biology of *Sleeping Beauty*

The *Sleeping Beauty* (SB in figure) transposon system is a two part system. It consists of the transposase enzyme and its transposon substrate. The *Sleeping Beauty* transposase recognizes inverted repeat/direct repeat (IR) elements that flank a given sequence of DNA. The transposase can excise the transposon from its original location leaving behind a canonical three base footprint (CA/TG). It can subsequently reinsert the sequence at a new location anywhere a TA dinucleotide is present. This TA is duplicated on reinsertion<sup>76</sup>. Increasing transposon size or the expression of transposase can reduce the amount of transposition<sup>97</sup>, whereas methylation of the transposon substrate seems

to increase it<sup>98</sup>. Host factors that have been linked to *Sleeping Beauty*-mediated transposition include the DNA-bending, high-mobility group protein B1 (HMGB1)<sup>99</sup>, and DNA repair proteins including the DNA-dependent protein kinase catalytic subunit (DNA-PKcs) and DNA-binding subunits (Ku70 and Ku86), the ataxia-telangiectasia mutated gene (*Atm*), and X-ray repair cross-complementing protein 4 (Xrcc4)<sup>100,101</sup>.



# ***Sleeping Beauty***

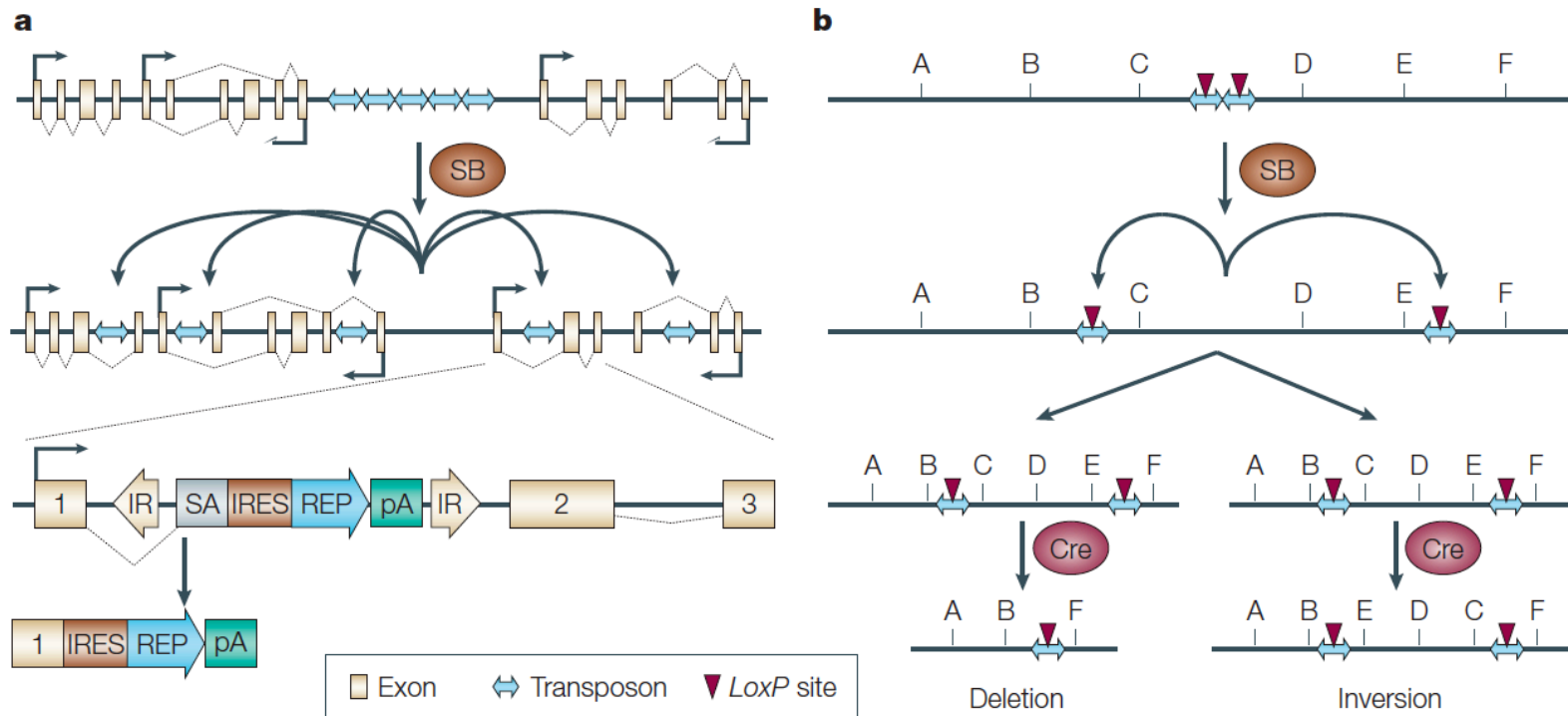


Figure 4 | **Exploiting the 'local hopping' of the *Sleeping Beauty* transposon.** Transposons that reside on chromosomes mobilize to loci that are linked to the donor locus approximately 50–80% of the time, with most of the loci reintegrating in close vicinity of the transgene insertion site. This 'local hopping' phenomenon can be used for region-specific mutagenesis and chromosome engineering. **a** | Mice doubly transgenic for a specific mutagenic transposon transgene and a *Sleeping Beauty* (SB) transposase transgene can be crossed with wild-type mice to generate offspring with heritable, novel transposon insertion events. On gene insertion, the gene-trap transposon can elicit premature truncation of endogenous transcripts. Transposons tend to insert at loci linked to the donor site, and so a single transgene could be used to mobilize transposons to saturate a given region of the genome with mutations. **b** | Alternatively, transposons containing *loxP* sites could be mobilized to produce unique chromosomal substrates for Cre-mediated recombination. Mice doubly transgenic for a *loxP*-containing transposon and transposase transgenes are bred to generate offspring with new transposon insertions at linked loci. The resulting animals now harbour two *loxP* sites on the same chromosome and can subsequently be crossed to a Cre-expressing strain to produce chromosomal rearrangements, including deletions and inversions. IRES, internal ribosomal entry site; IR, inverted repeat/direct repeat; pA, polyadenylation signal; REP, reporter; SA, splice acceptor.

**Table 1. International Mouse Gene Knockout Programs**

Type of Resource	Type of Knockout	2006	2007	2008	2009	2010	Totals
KOMP							
ES Cell	Targeted Deletion	175	500	941	942	942	3500
Mouse	Targeted Deletion	50	50	50	50	50	250
ES Cell	Targeted Conditional	1000	1000	1000	1000	1000	5000
Mouse	Targeted Conditional	50	50	50	50	50	250
EUCOMM							
ES Cell	Trapped Conditional	3000	6000	3000			12,000
ES Cell	Targeted Conditional	1000	3000	4000			8000
Mouse	Mixed	20	100	200			320
NorCOMM							
ES Cell	Trapped Conditional	1000	4000	3000	2000		10,000
ES Cell	Targeted Conditional	100	400	750	750		2000
Mouse	Mixed	25	25	25	25		100
Cumulative for All Programs							
ES Cell	Trapped (Conditional)	4000	14000	20000	22000		22,000 <sup>a</sup>
ES Cell	Targeted (Deletion)	175	675	1616	2558	3500	3500
ES Cell	Targeted (Conditional)	2100	6500	12,250	14,000	15,000	15,000
Mouse	Mixed	125	400	715	830	920	920

NorCOMM (North American Conditional Mouse Mutagenesis Project, <http://norcomm.phenogenomics.ca/index.htm>)

EUCOMM (EUropean Conditional Mouse Mutagenesis Program, <http://www.eucomm.org>)

KOMP (KnockOut Mouse Project, <http://www.knockoutmouse.org>)

## **Additional gene trap databases**

<http://www.tigm.org>

<http://www.genetrap.org>

<http://www.lexicon-genetics.com/discovery/omnibank.htm>

<http://www.deltagen.com>

<http://www.sanger.ac.uk/PostGenomics/mousegenomics/>

## **Phenotype driven approaches**

# **Relationship between the number of components encoded in the mammalian genome and the complexity of the phenome**

Yeast ( <i>Saccharomyces cerevisiae</i> )	~6000 genes
Worm ( <i>Caenorhabditis elegans</i> )	~19,000 genes
Fruitfly ( <i>Drosophila melanogaster</i> )	~14,000 genes
Man ( <i>Homo sapiens</i> )	22,000–28,000 genes

**How is the large increase in phenomic diversity encoded by a genome with only a modest increase in the gene count?**

# Relationship between the number of components encoded in the mammalian genome and the complexity of the phenome

Yeast ( <i>Saccharomyces cerevisiae</i> )	~6000 genes
Worm ( <i>Caenorhabditis elegans</i> )	~19,000 genes
Fruitfly ( <i>Drosophila melanogaster</i> )	~14,000 genes
Man ( <i>Homo sapiens</i> )	22,000–28,000 genes

**How is the large increase in phenomic diversity encoded by a genome with only a modest increase in the gene count?**

Beadle/Tatum/Garrod: “one gene - one protein - one trait”

From a biochemical perspective, however, specific traits are rarely encoded by a single protein but are usually produced by assemblies of many proteins working in concert!



The magnitude of the potential increase in phenomic diversity created by combinatorial assembly and a doubling of the gene count can be estimated by making the following conservative assumptions:

(a) the human mouse genomes comprise 28,000 genes encoding 28,000 proteins

(b) the Drosophila genome comprises ~14,000 genes and proteins

(c) the average multiprotein assembly specifying a particular biochemical trait comprises products of 10 different genes

Human:  $28,000!/(10! \ 27,990!) = 8.15 \times 10^{37}$

Drosophila:  $14,000!/(10! \ 13,990!) = 7.95 \times 10^{34}$

A doubling of gene/protein numbers encoded in the genome yields a 1000-fold increase in possible 10-mer protein combinations available to specify the phenome.

If one assumes that 10 subcellular traits (each specified by a set of 10 proteins) combine to specify a cellular trait, then the doubling of the protein set in mammals compared with flies yields an increase in the number of potential cellular traits by a factor of  $10^{30}$  (from  $10^{333}$  to  $10^{363}$ )

**Table 2** Number of unique subcellular protein combinations that can potentially be formed from a genome encoding 7000 (e.g., yeast), 14,000 (e.g., fruitfly), 28,000 (e.g., human/mouse), 56,000, and 112,000 genes/proteins

Genes/proteins per genome	Number of unique subcellular protein combinations possible <sup>a</sup>			Number of cellular traits (combinations of 10-protein combinations) <sup>b</sup>
	10-protein combinations	20-protein combinations	40-protein combinations	
7000	$7.73 \times 10^{31}$	$3.19 \times 10^{58}$	$6.98 \times 10^{105}$	$\sim 10^{303}$
14,000	$7.95 \times 10^{34}$	$3.39 \times 10^{64}$	$8.11 \times 10^{117}$	$\sim 10^{333}$
28,000	$8.15 \times 10^{37}$	$3.58 \times 10^{70}$	$9.17 \times 10^{129}$	$\sim 10^{363}$
56,000	$8.35 \times 10^{40}$	$3.77 \times 10^{76}$	$1.02 \times 10^{142}$	$\sim 10^{393}$
112,000	$8.56 \times 10^{43}$	$3.96 \times 10^{82}$	$1.13 \times 10^{154}$	$\sim 10^{423}$

<sup>a</sup>Calculations assume the average subcellular biochemical trait is specified by combinations of 10, 20, or 40 different proteins.

<sup>b</sup>Cellular traits are assumed to be specified by combinations of ten different subcellular traits, each specified by a 10-protein combination.

Systematic efforts to knock out or knock down all of the mammalian genes in mice starts with the assumption of “one gene - one protein - one trait.”

However, most of the genome-phenome code will only be revealed by analyzing genetic variants that interfere with specific protein-protein interactions, notably point mutations arising by natural variation or induced experimentally by chemicals such as ENU.

# N-ethyl-N-nitrosourea (ENU) induced mutants



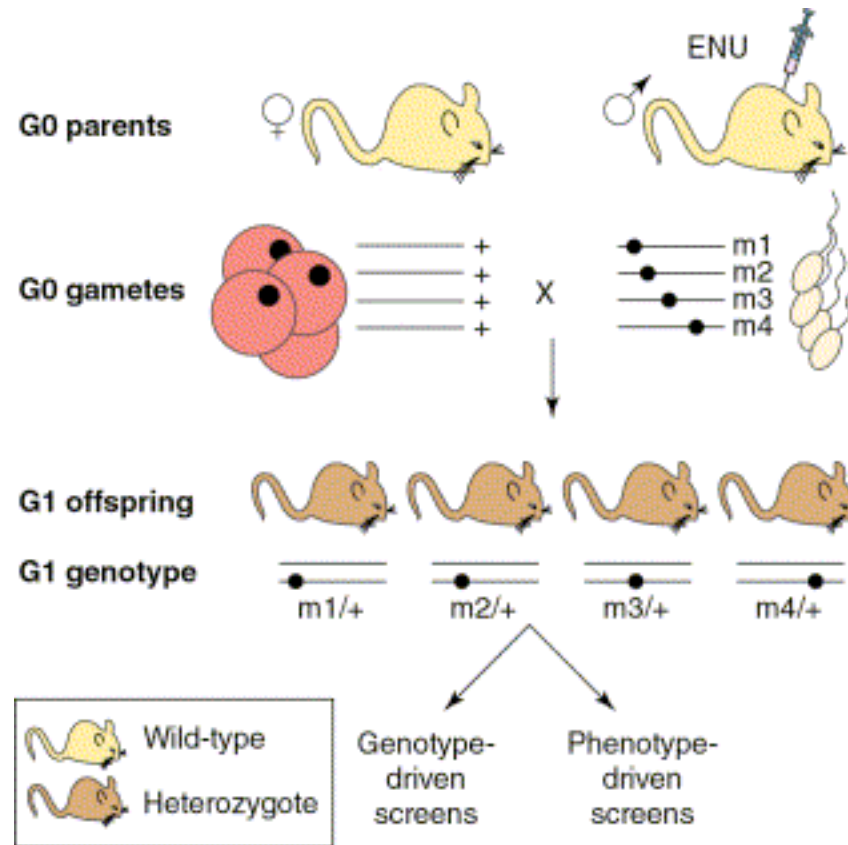
ENU induces one new loss-of-function mutation per gene in every ~700 gametes of treated male mice and one nucleotide change in every one megabase (Mb) of genomic DNA

The most common mutations induced by ENU are  $A/T \rightarrow T/A$  transversions and  $A/T \rightarrow G/C$  transitions

When translated into a protein product, these changes result in 64% missense mutations, 10% nonsense, and 26% cause errors in mRNA splicing

Genetic variation induced in mice by ENU is thus ideally suited to reveal discrete actions of proteins working combinatorially by selectively altering individual protein domains and splicing products in the same way as natural variation, while at the same time inducing a very low frequency of linked irrelevant sequence changes.

# Ethylnitrosourea (ENU) induced mutants



*Drug Discovery Today*

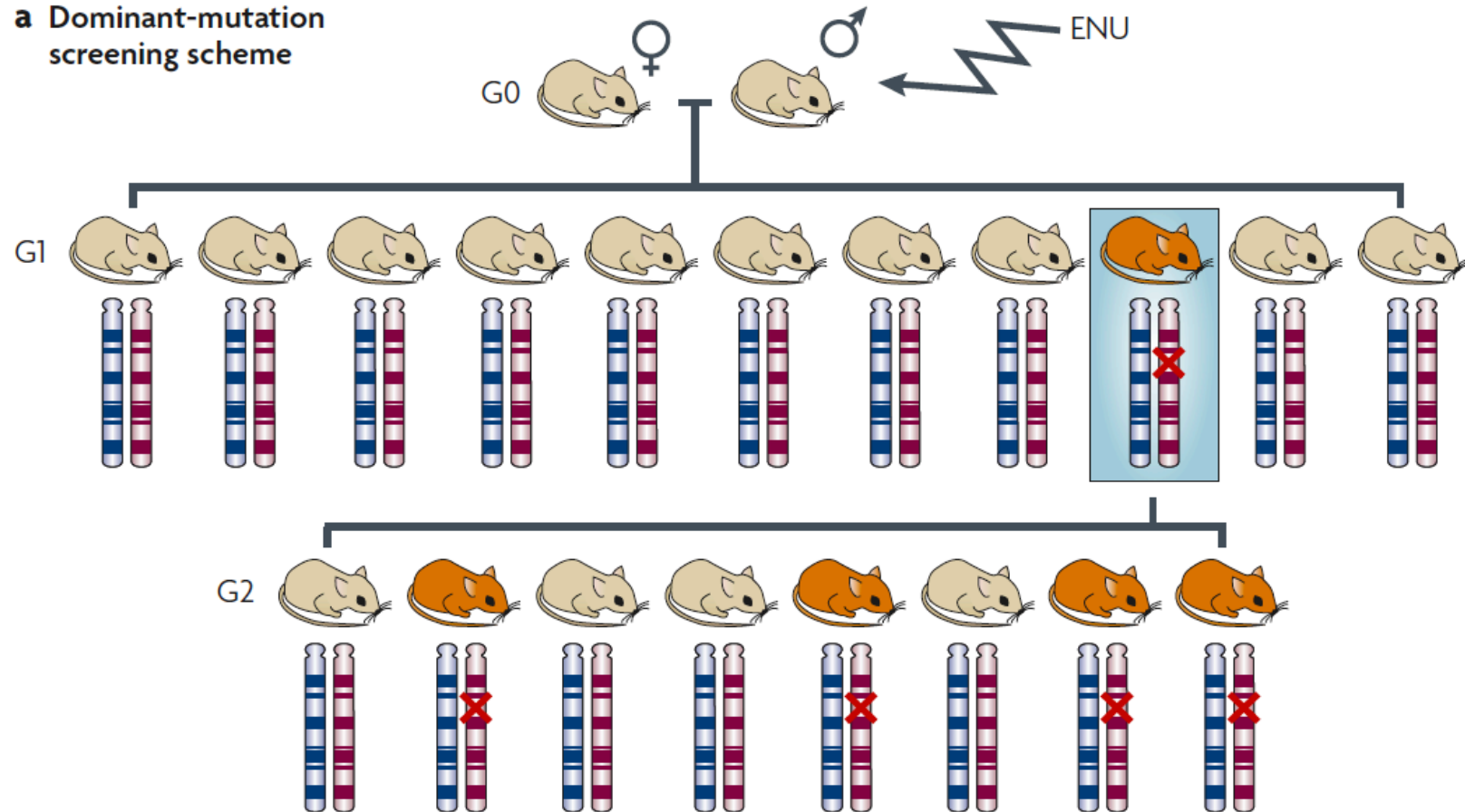
[http://www.cmhd.ca/enu\\_mutagenesis/index.html](http://www.cmhd.ca/enu_mutagenesis/index.html)

<http://www.gsc.riken.go.jp/Mouse/>

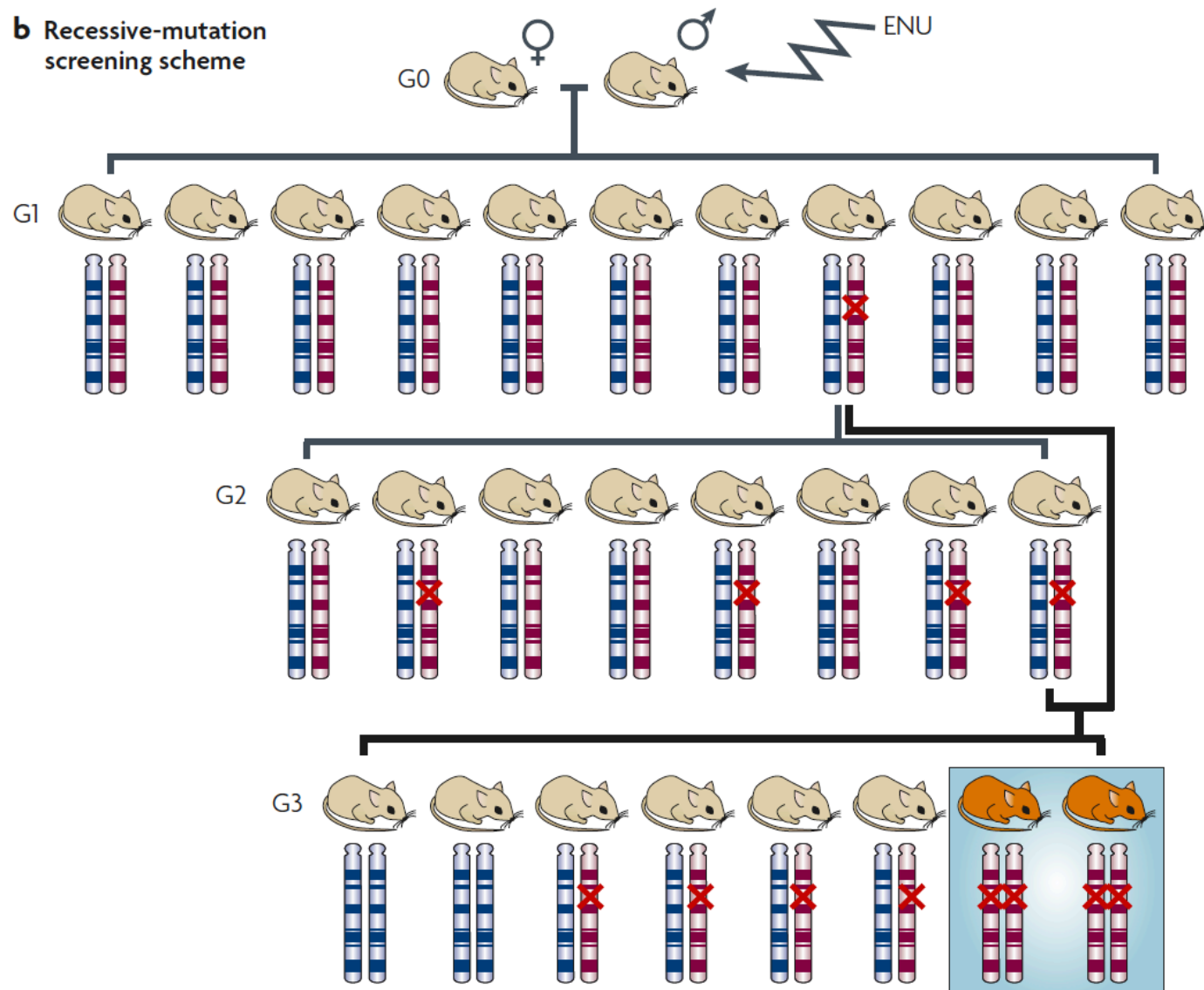
<http://www.ingenium-ag.com/>

# Genome-wide dominant and recessive protocols for phenotype-driven mutagenesis screens

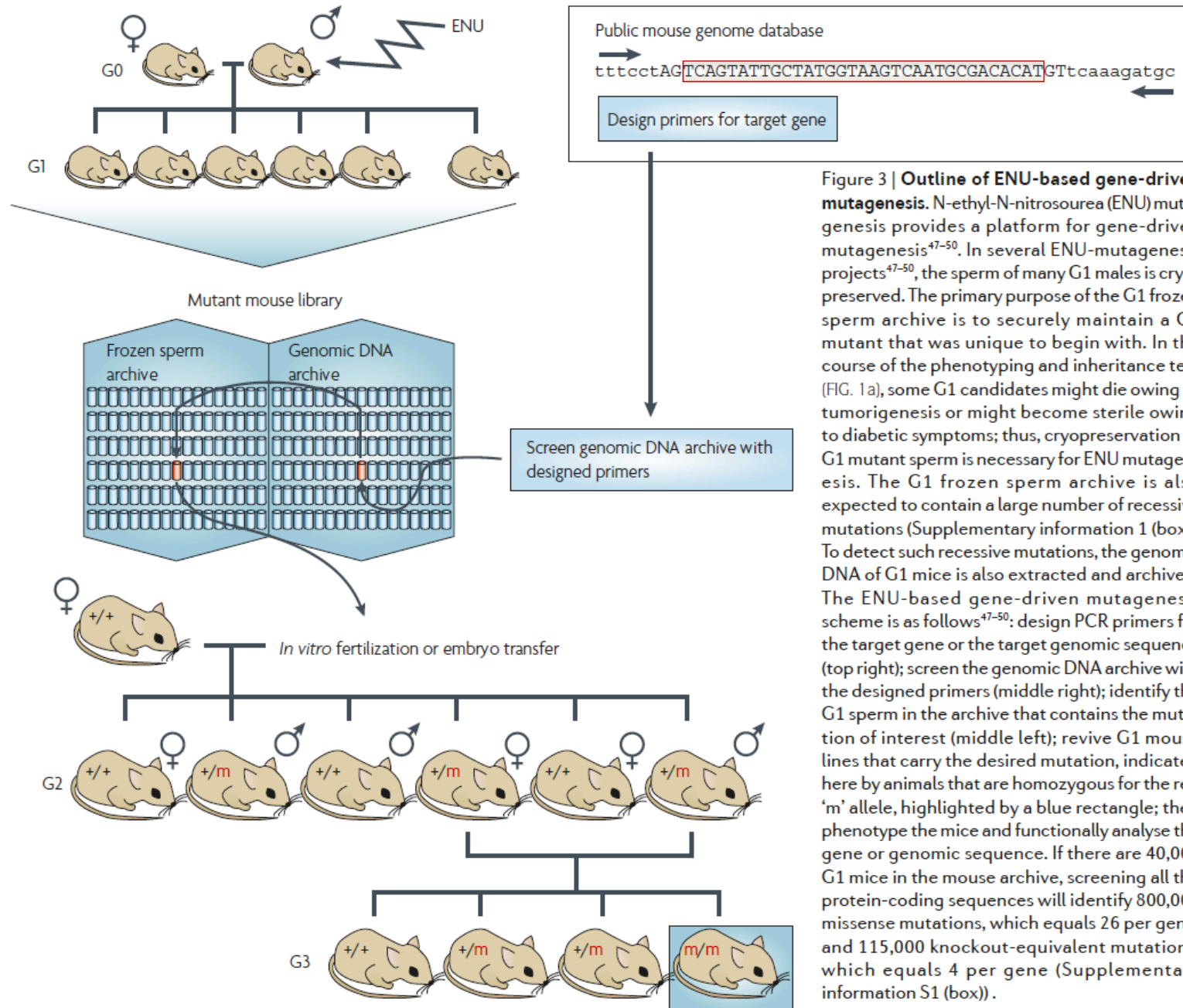
**a** Dominant-mutation screening scheme



# Genome-wide dominant and recessive protocols for phenotype-driven mutagenesis screens



# ENU-based **gene-driven** mutagenesis



**Figure 3 | Outline of ENU-based gene-driven mutagenesis.** N-ethyl-N-nitrosourea (ENU) mutagenesis provides a platform for gene-driven mutagenesis<sup>47–50</sup>. In several ENU-mutagenesis projects<sup>47–50</sup>, the sperm of many G1 males is cryopreserved. The primary purpose of the G1 frozen sperm archive is to securely maintain a G1 mutant that was unique to begin with. In the course of the phenotyping and inheritance test (FIG. 1a), some G1 candidates might die owing to tumorigenesis or might become sterile owing to diabetic symptoms; thus, cryopreservation of G1 mutant sperm is necessary for ENU mutagenesis. The G1 frozen sperm archive is also expected to contain a large number of recessive mutations (Supplementary information 1 (box)). To detect such recessive mutations, the genomic DNA of G1 mice is also extracted and archived. The ENU-based gene-driven mutagenesis scheme is as follows<sup>47–50</sup>: design PCR primers for the target gene or the target genomic sequence (top right); screen the genomic DNA archive with the designed primers (middle right); identify the G1 sperm in the archive that contains the mutation of interest (middle left); revive G1 mouse lines that carry the desired mutation, indicated here by animals that are homozygous for the red 'm' allele, highlighted by a blue rectangle; then phenotype the mice and functionally analyse the gene or genomic sequence. If there are 40,000 G1 mice in the mouse archive, screening all the protein-coding sequences will identify 800,000 missense mutations, which equals 26 per gene, and 115,000 knockout-equivalent mutations, which equals 4 per gene (Supplementary information S1 (box)).



If we assume that a target gene consists of 2,000 base pairs of protein-coding sequences then, on average, already 80 point mutations exist per gene in the frozen sperm archives.

Possibility to generate allelic series for most or all genes.

# Congenic mice

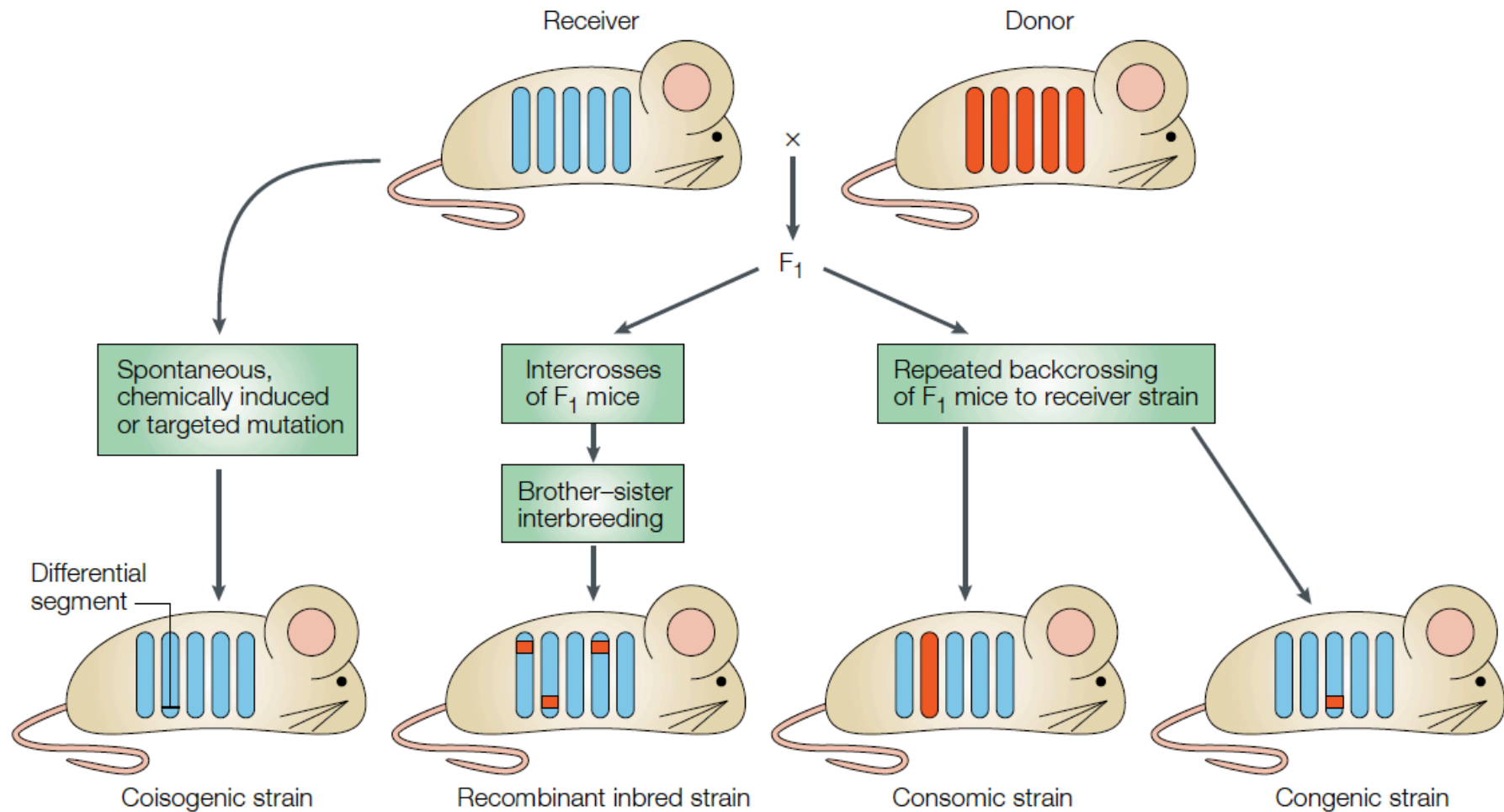
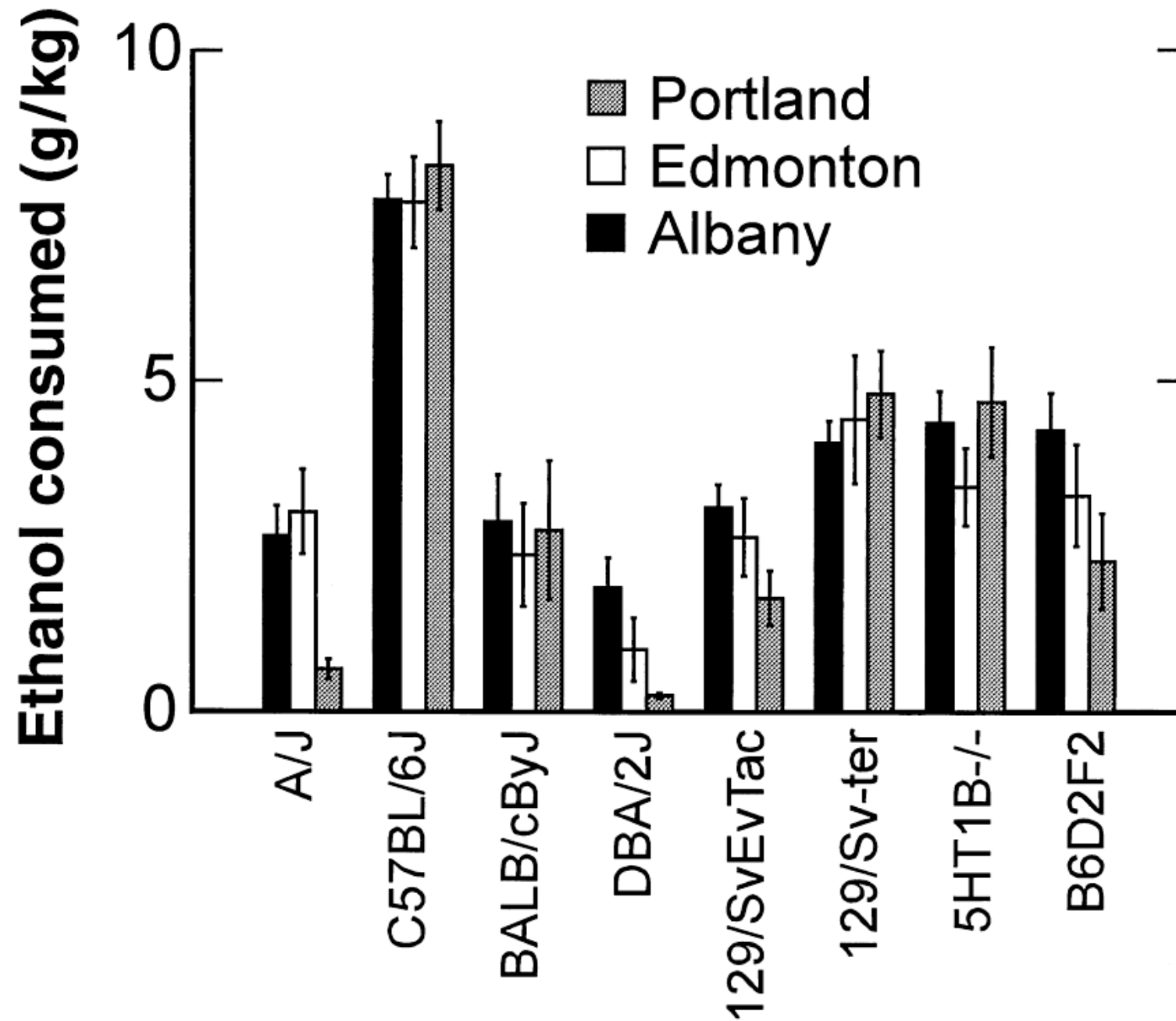
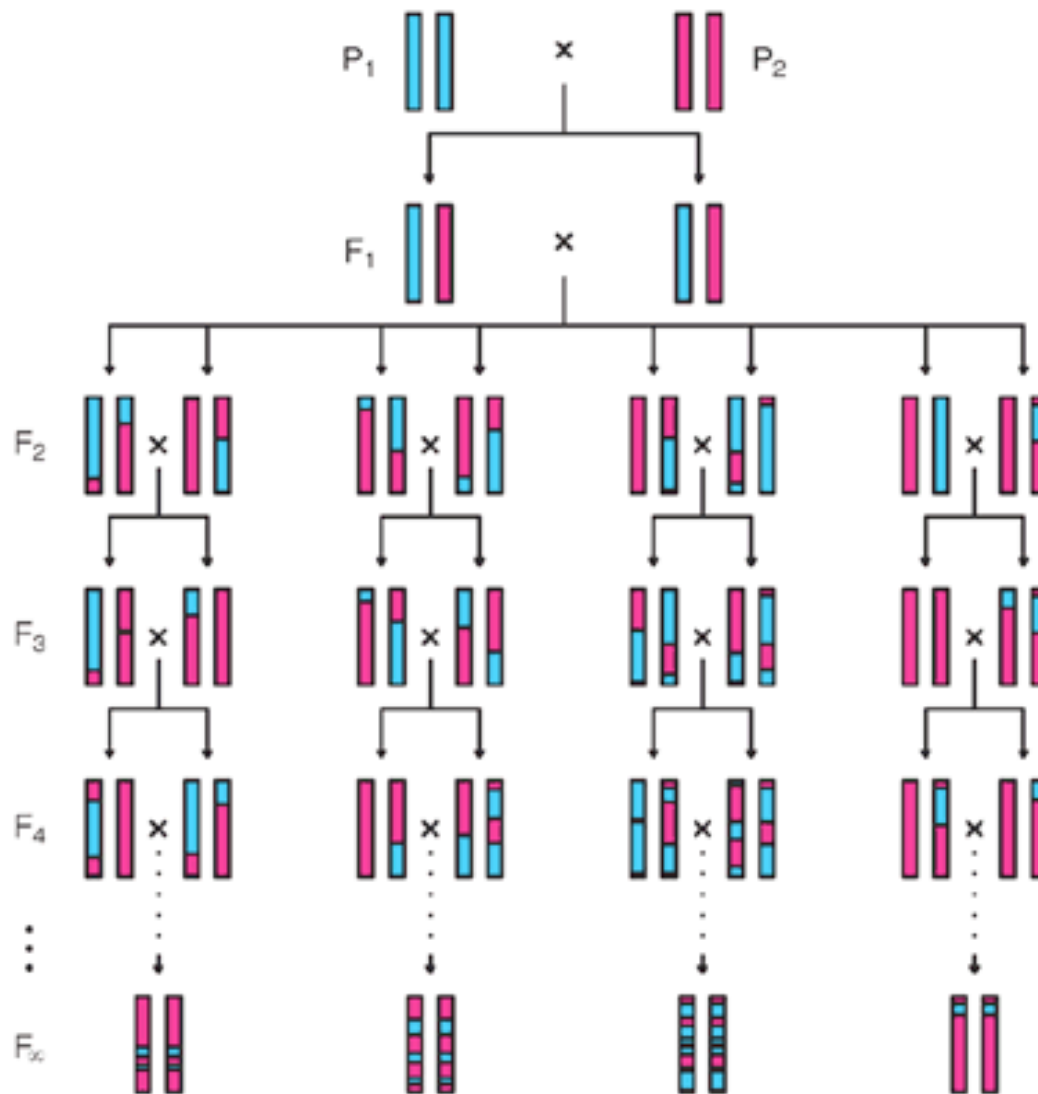


Figure 1 | **Selected analytical tools in mouse genetics.** The generation of coisogenic, recombinant inbred, consomic and congenic strains is illustrated.



# Recombinant inbred (RI) strains



*Nature Genetics* 36, 1133 - 1137 (2004)

*Nature Genetics* 37, 209 - 210 (2005)

## Recombinant inbred (RI) strains

**BXD:** derived by crossing C57BL/6J (B6) and DBA/2J (D2); ~90 strains

**AXB/BXA:** derived by crossing A/J (A) and C57BL/6J (B); ~25 strains

**BXH:** derived by crossing C57BL/6J (B) and C3H/HeJ (H); ~16 strains

**CXB:** derived by crossing BALB/cBy and C57BL/6By; ~13 strains

**AKXD:** derived by crossing AKR/J (AK) and DBA/2J (D); ? Strains

**LXS:** derived by crossing Inbred Long-Sleep (ILS) and Inbred Short-Sleep (ISS) strains; 77 strains

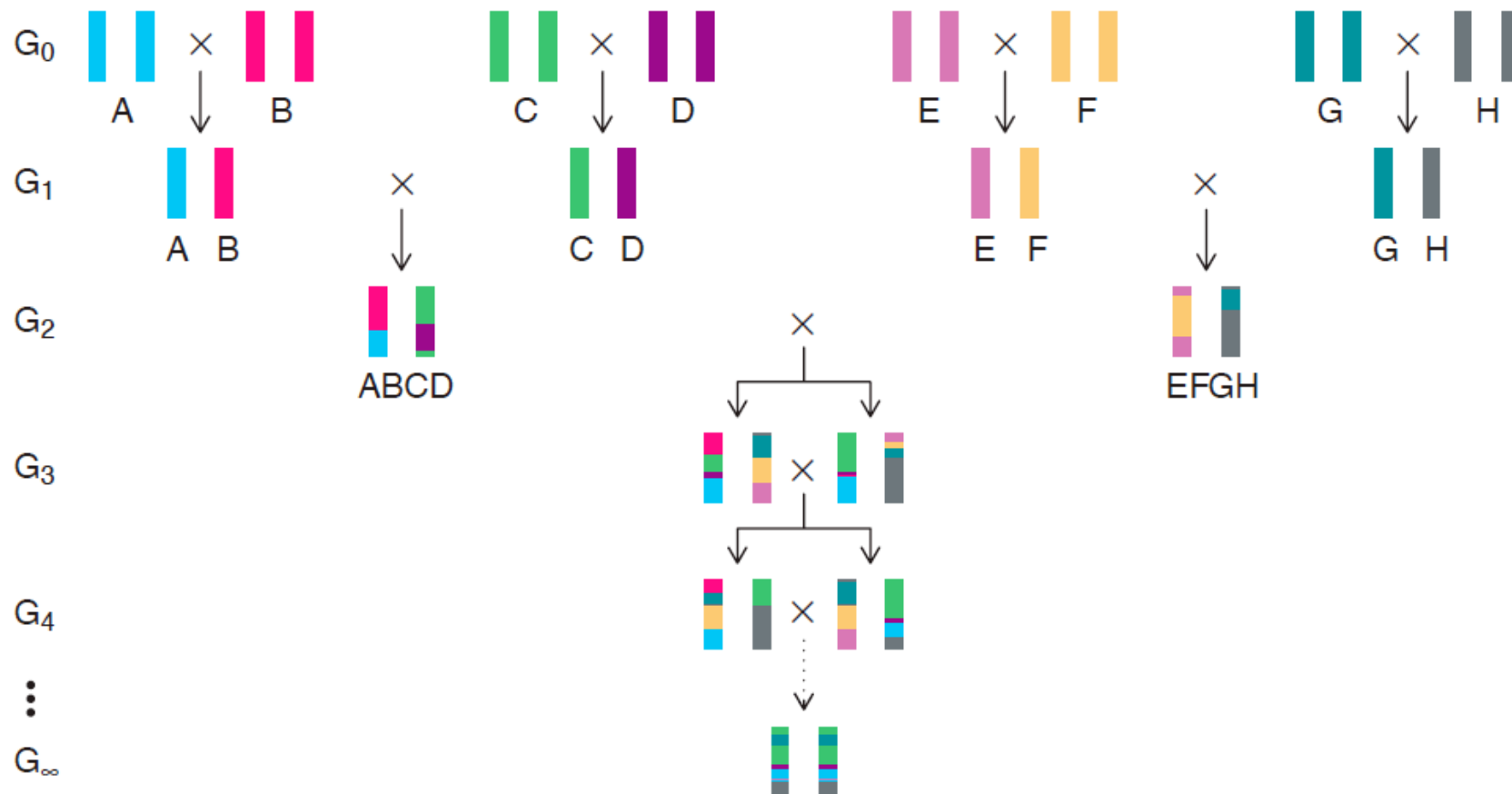
<http://www.well.ox.ac.uk/mouse/INBREDS/RIL/BXD.shtml>

<http://www.genenetwork.org/mouseCross.html>

# The Collaborative Cross, a community resource for the genetic analysis of complex traits

The Complex Trait Consortium\*

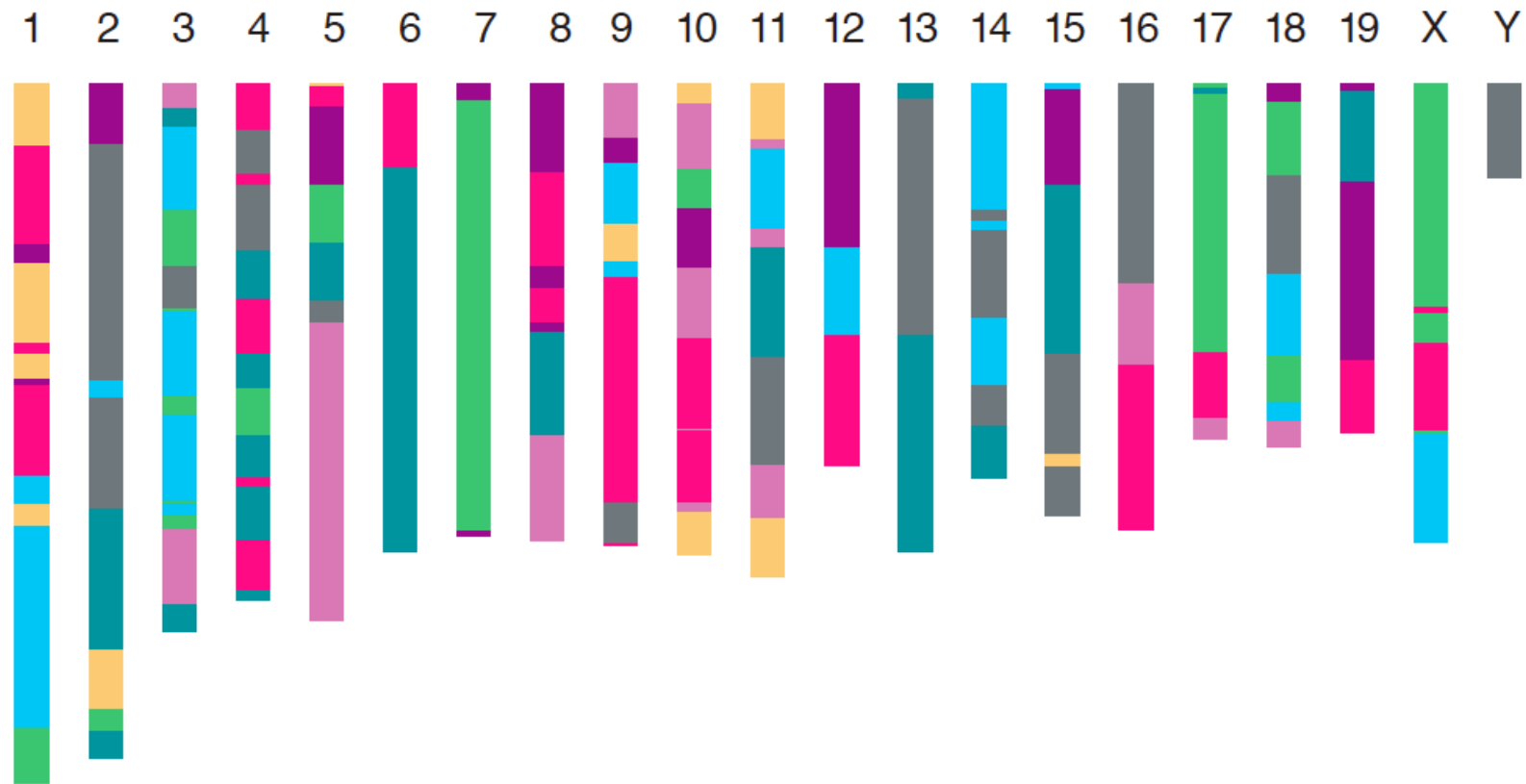
The goal of the Complex Trait Consortium is to promote the development of resources that can be used to understand, treat and ultimately prevent pervasive human diseases. Existing and proposed mouse resources that are optimized to study the actions of isolated genetic loci on a fixed background are less effective for studying intact polygenic networks and interactions among genes, environments, pathogens and other factors. The Collaborative Cross will provide a common reference panel specifically designed for the integrative analysis of complex systems and will change the way we approach human health and disease.



**Figure 2** The eight-way 'funnel' breeding scheme for generating RI strains. G, generation.

8 inbred strains: A/J, C57BL/6J, 129S1/SvImJ, NOD/LtJ, NZO, CAST/Ei, PWK/Ph, WSB/E

All 8 genomes are brought together in G<sub>2</sub>:F<sub>1</sub>, and the offspring of this cross are inbred. 90% inbreeding is expected at G<sub>2</sub>:F<sub>20</sub> based on theoretical results.



**Figure 3** A typical eight-way RI strain. Color scheme indicates the parental origin of genomic segments.

Aim is to generate 1,000 Collaborative Cross (RI) strains



# Phenotyping mouse models

Considering the 20,000 to 25,000 mammalian genes, each with one or more mutant alleles, and the number of primary phenotypic parameters that can be measured (currently between 300 and 400 parameters per mouse line) in cohorts of male and female mice, the question arises of how an endeavour of such a scope may be accomplished.

$25'000 \times 400 = 10'000'000$  measures

Mouse Phenome Project / Mouse Phenome Database

<http://phenome.jax.org/pub-cgi/phenome/mpdcgi?rtn=docs/home>

Phenotype screens	Methods	Age of mice (weeks)											
		8	9	10	11	12	13	14	15	16	17	18	
Pipeline 1													
Dysmorphology	Anatomical observation		•										
	DEXA, X-ray							•					
Cardiovascular	Blood pressure				•								
	Heart weight								•				
Energy metabolism	Calorimetry					•							
Clinical chemistry	Simplified IPGTT						•						
Eye	Eye size (LIB)								•				
Lung function	Plethysmography									•			
Molecular phenotyping	Expression profiling									•			
Pipeline 2													
Behaviour	Open field		•										
	Acoustic startle and PPI				•								
Neurology	Modified SHIRPA, grip strength, rotarod		•	•									
Nociception	Hot plate					•							
Eye	Ophthalmoscopy and slit lamp						•						
Clinical chemistry	Clinical chemical analysis, haematology							•				•	
Immunology	FACS analysis of PBCs, immunoglobulin concentration							•				•	
Steroid metabolism	DHEA, testosterone							•				•	
Cardiovascular	ANP, ECG or echocardiogram							•	•	•	•		
Pathology	Macro and microscope analysis												•

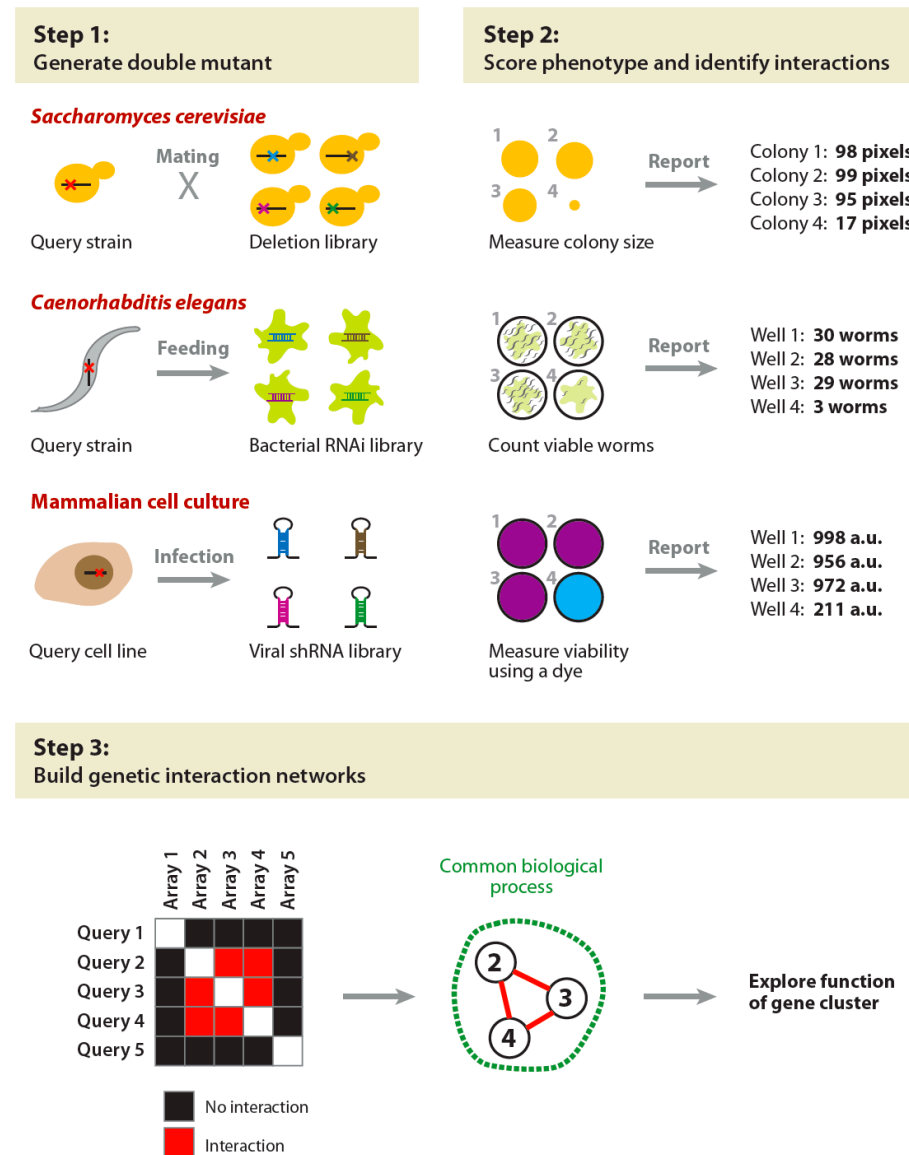
**Figure 2 | Scheme of the primary phenotyping protocol of the German Mouse Clinic (GMC).** This scheme includes the EMPReSS slim primary phenotyping protocol, which is a common standard of European mouse clinics. Screens such as molecular phenotyping, lung function, steroid metabolism and pathological screens are performed in addition to the EMPReSS slim protocol. The GMC primary phenotyping screen starts 2 weeks after the mutant mouse lines are imported at the age of 9 weeks. For phenotypic analysis, the mice are distributed into one of two pipelines, in which they are subjected to a defined series of tests. The primary screen ends at the age of 18 weeks. Based on the results of the screens, decisions for secondary and tertiary screens are made. ANP, atrial natriuretic peptide; DEXA, dual-energy X-ray absorption; DHEA, dehydroepiandrosterone; ECG, electrocardiogram; FACS, fluorescence-activated cell sorting; IPGTT, intraperitoneal glucose tolerance test; LIB, laser interference biometry; PBC, peripheral blood cell; PPI, pre-pulse inhibition; SHIRPA, a protocol for comprehensive behaviour assessment. Figure is modified, with permission, from REF. 111 © Humana Press (2009).

# Archiving and dissemination of mouse models

Money!

1 cage (max 5 mice) = 15CHF/month

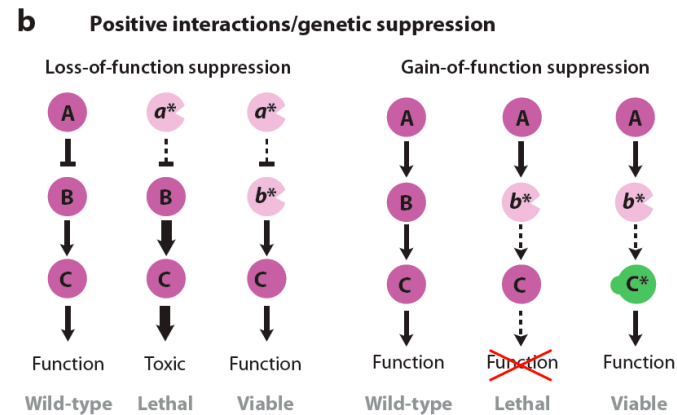
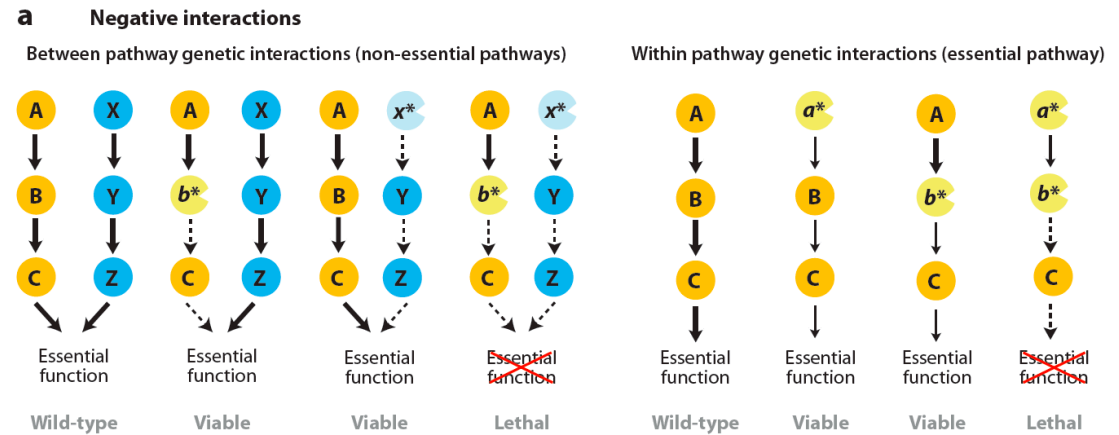
# Identifying genetic interactions in a systematic manner



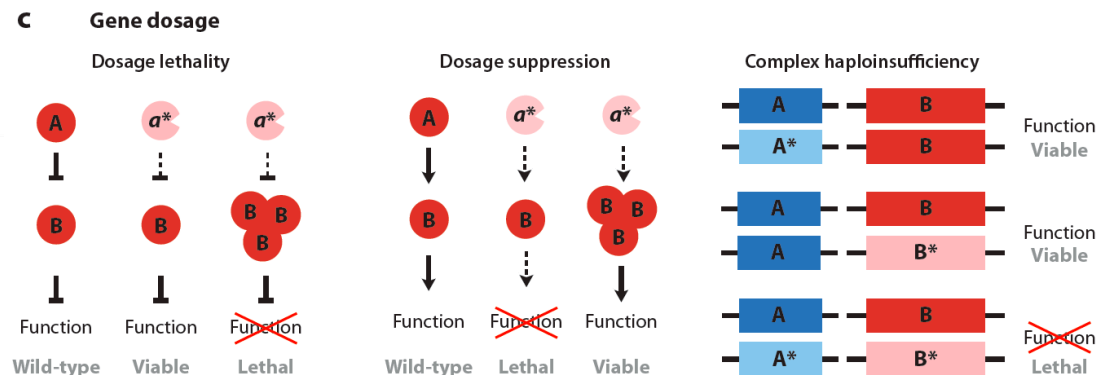
**Figure 1**

An outline of systematic screening strategies in three prominent model systems, *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, and mammalian cell culture. All systematic screens follow a similar pattern involving the generation of double mutants (*Step 1*), the scoring of a double mutant phenotype, which must be compared with the corresponding single mutant phenotypes (*Step 2*) and the construction and interpretation of the resulting genetic interaction data (*Step 3*). In *Step 1*, the black line with a red x indicates a mutated gene in a chromosome. In *Step 2*, a.u. stands for arbitrary units.

# A representation of the molecular mechanisms underlying different classes of genetic interactions



A representation of the molecular mechanisms underlying different classes of genetic interactions. (a) Negative interactions can arise from the disruption of parallel pathways converging on a common process (between pathway genetic interactions) or by decreasing the flux through the same essential pathway (within pathway genetic interactions). (b) Positive interactions/genetic suppression. Mutation of a negative regulator ( $a^*$ ) leads to hyperactivation of the pathway and accumulation of a toxic gene product (C). Subsequent loss-of-function mutation of a downstream pathway component ( $b^*$ ) reduces flux through the pathway thereby suppressing the toxic effects caused by mutant  $a^*$ . Gain-of-function suppression may arise when a downstream or terminal pathway component acquires a mutation ( $C^*$ ) such that it is no longer dependent on upstream activation events. (c) Gene dosage. Increasing the dosage of a gene (B) can be lethal (dosage lethality) in the presence of a mutation in another gene ( $a^*$ ) when A negatively regulates the activity of B. The lethal effects of a mutated essential gene (A) can be suppressed by overexpression of a downstream pathway component (dosage suppression). Heterozygous mutation of two independent loci can result in complex haploinsufficient phenotypes.



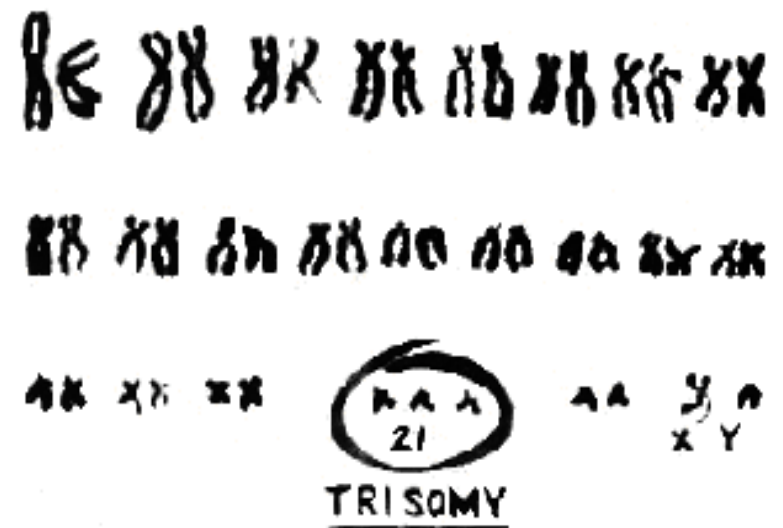
## Down syndrome



# Clinical phenotypes of Down syndrome

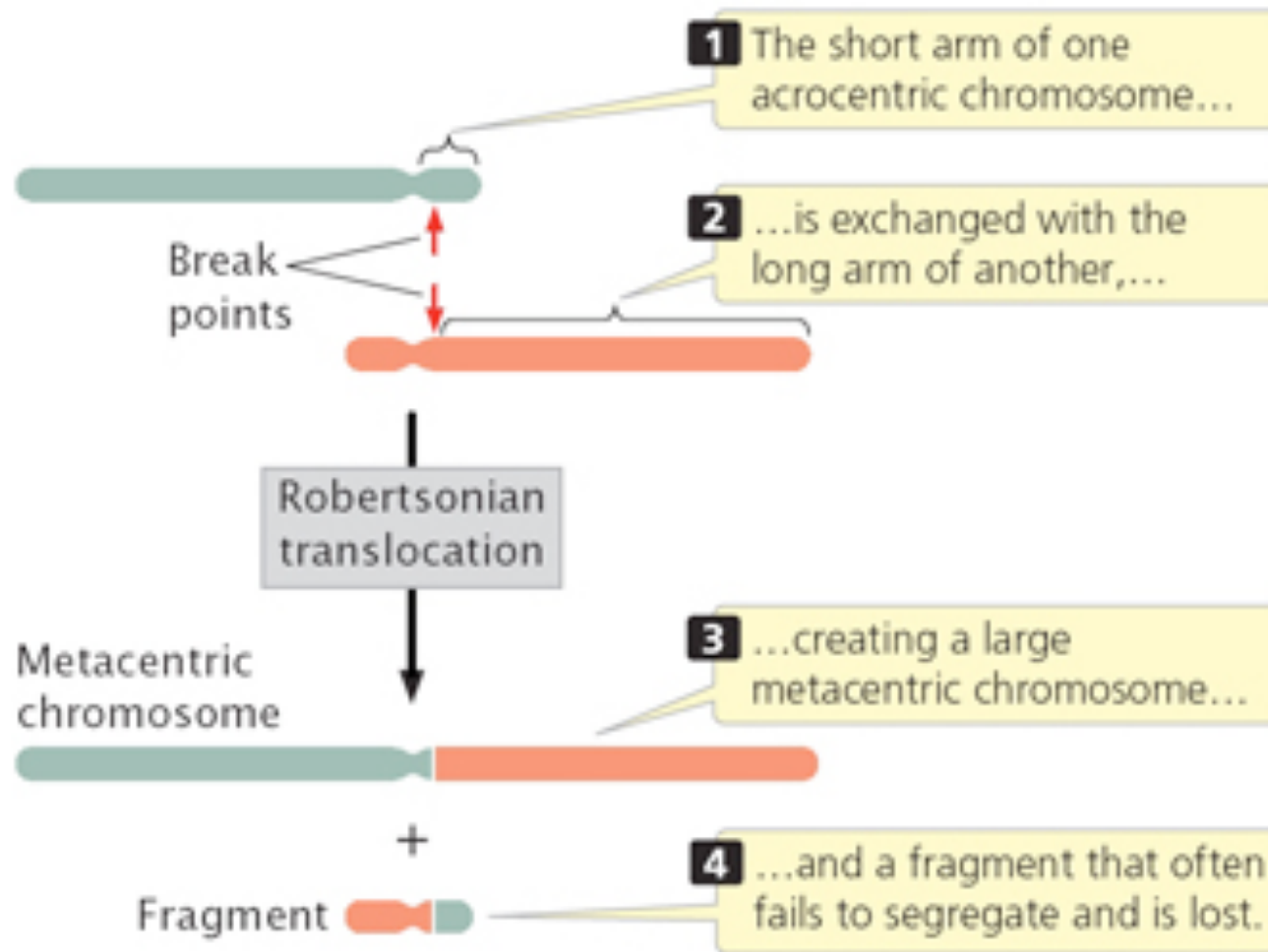
Phenotype	%	Phenotype	%
• Neurological :		♦ Limbs :	
– Mental retardation	100	– Short, broad hands	65
– Alzheimer disease	100 over 35 y.o.	– Short 5th finger	60
• Muscle :		♦ Skin :	
– Hypotonia	100	– Characteristic dermatoglyphics	85
• Growth :		♦ Cardiac :	
– Short stature	70	– Congenital heart defect	40
• Head :		– Atrioventricular canal	16
– Brachycephaly	75	♦ Gastrointestinal abnormalities :	
• Eyes :		– Duodenal stenosis/atresia	250x
– Epicanthic folds	60	– Imperforate anus	50x
– Iris Brushfield spots	55	– Hirschsprung disease	30x
• Mouth :		♦ Heme :	
– Protruding tongue	45	– Acute megakaryocytic leukemia	200-400x
• Ears :		– Leukemia (both ALL and AML)	10-20x
– Folded/dysplastic ear	50		

# Trisomy 21 or Down syndrome





# Robertsonian translocation



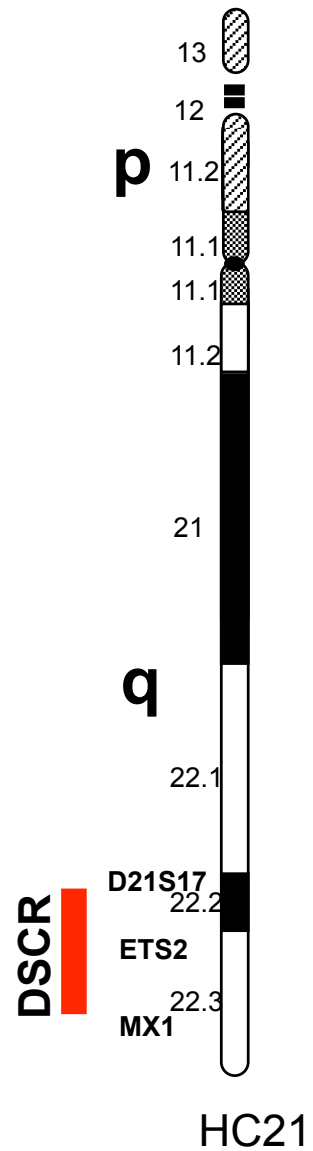
A person with a translocation is phenotypically normal. During reproduction, normal disjunctions leading to gametes have a significant chance of creating a gamete with an extra chromosome 21, producing a child with Down syndrome.

# Chromosomal abnormalities in Down syndrome

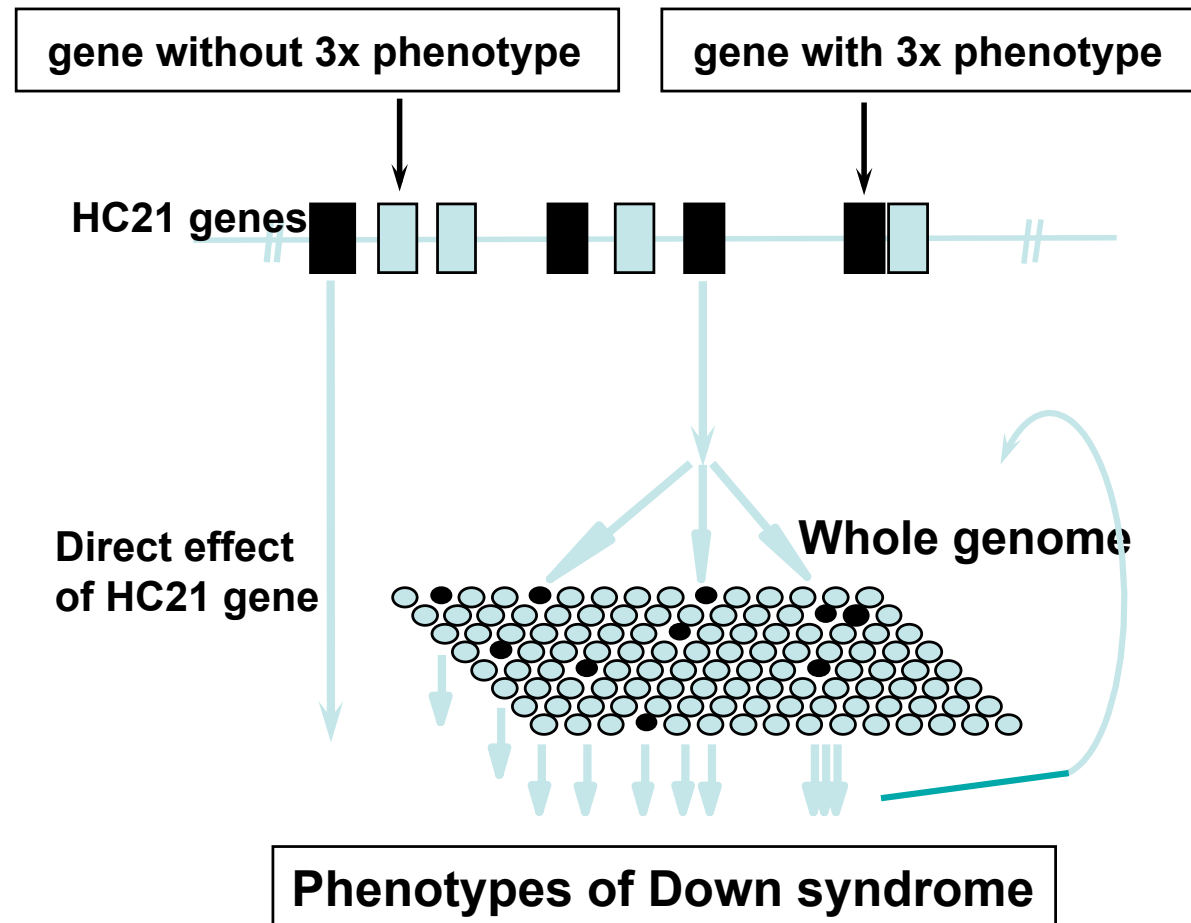
- ◆ Most common chromosomal abnormality
- ◆ 1 in 700 live births



# Down syndrome critical region (DSCR)



# Model for the pathogenesis of Down syndrome



# Proteins potentially involved in DS phenotypes

- Subunits of multimeric proteins
- Transcription factors, gene expression regulators
- Proteins involved in cell-cell interactions
- Receptors / Ligands / Signal transduction
- Pattern formation / Development
- Enzymes (proteases, deaminases, ...)

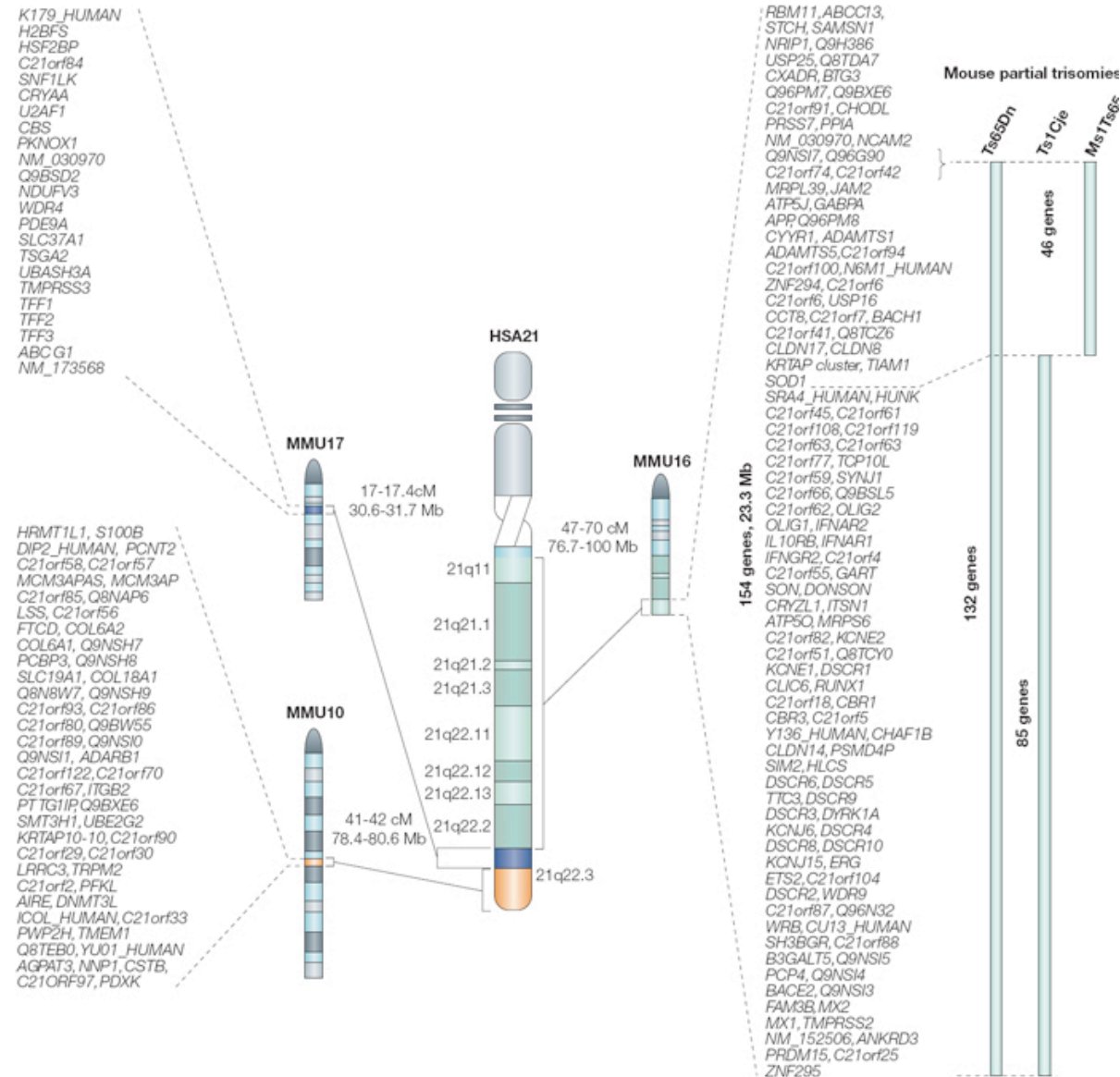
..... so pretty much every protein is a good candidate!

# How to model Down syndrome?

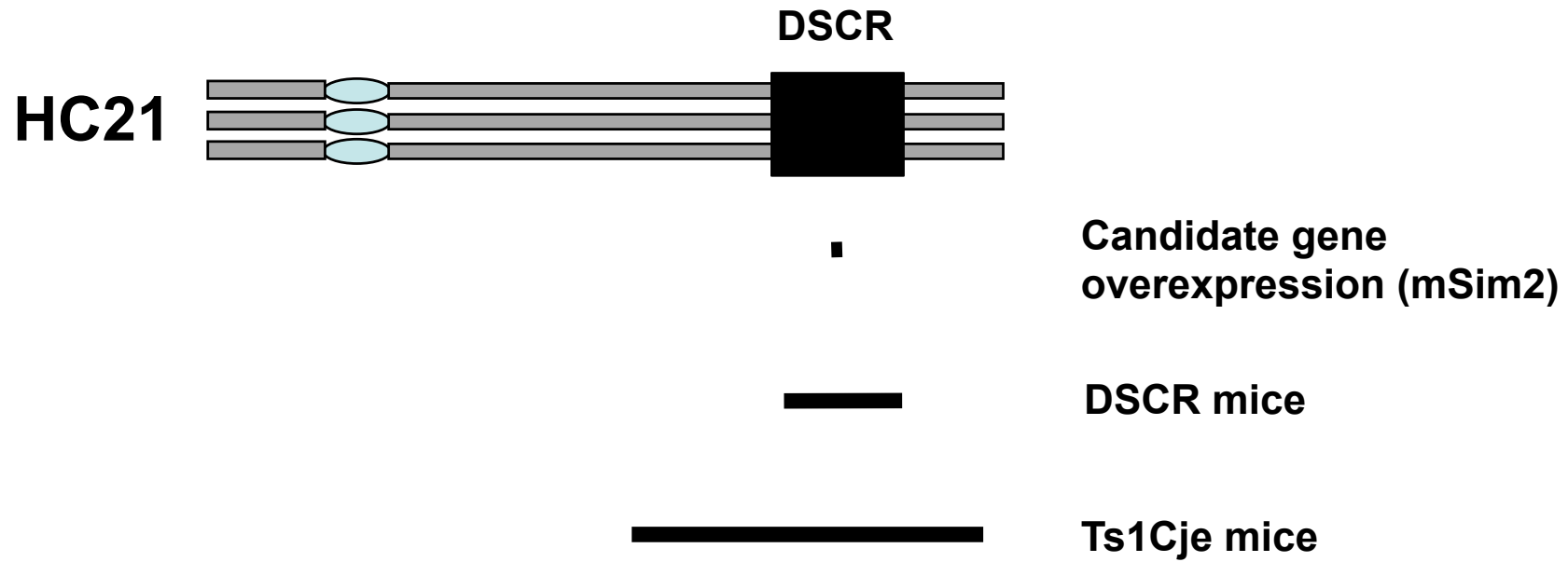


Human - mouse synteny issue

# Regions of synteny between human chromosome 21 (HSA21) and mouse chromosomes (MMUs) 16, 17, and 10.

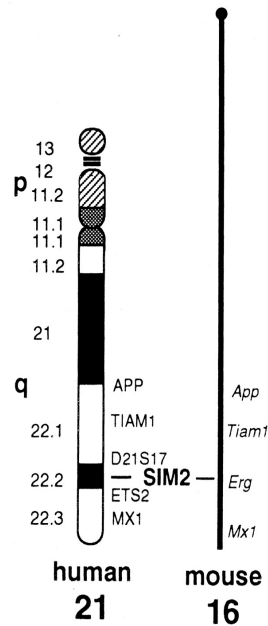


# How to model Down syndrome?



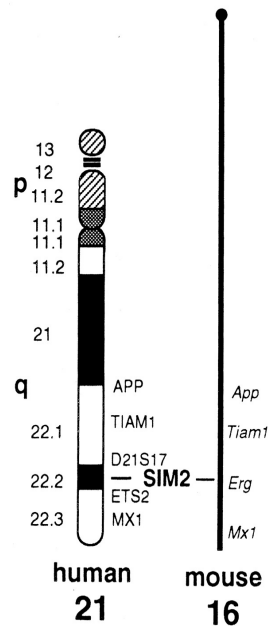


# SIM2 as a candidate for some of the DS Phenotypes

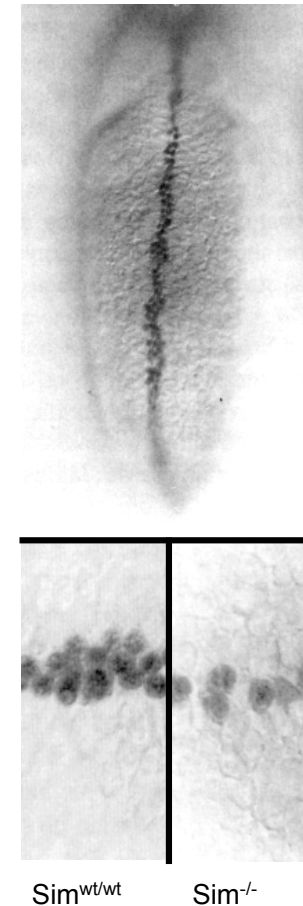


- ◆ The mapping of hSIM2 on HC21 in the so-called DS critical region

# SIM2 as a candidate for some of the DS Phenotypes

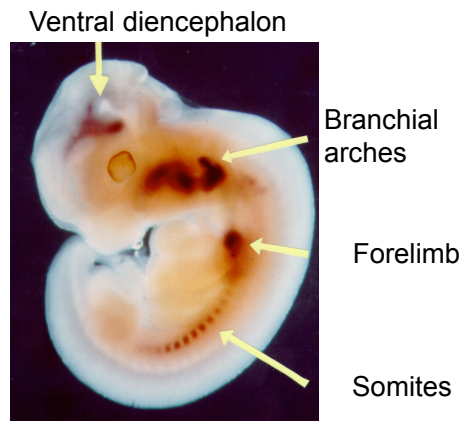
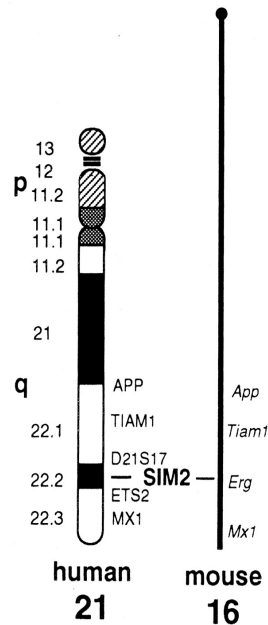


- ◆ The mapping of hSIM2 on HC21 in the so-called DS critical region
- ◆ Its Drosophila homologue (dsim) is a master regulator of fruitfly neurogenesis expressed early in embryogenesis in the developing midline



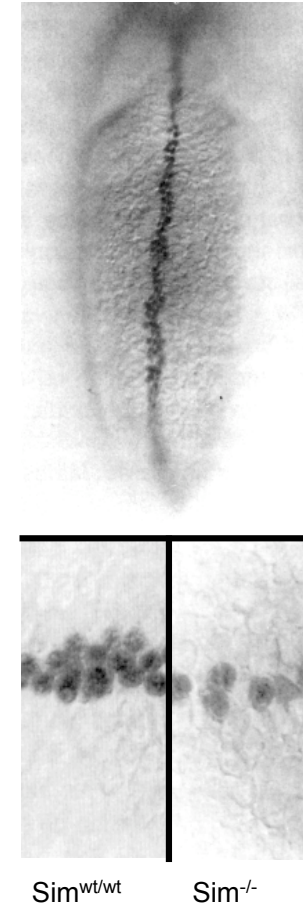
Nambu et al., 1991, Cell 67: 1157-1167.  
Crews et al., 1988, Cell 52: 143-151.

# SIM2 as a candidate for some of the DS Phenotypes



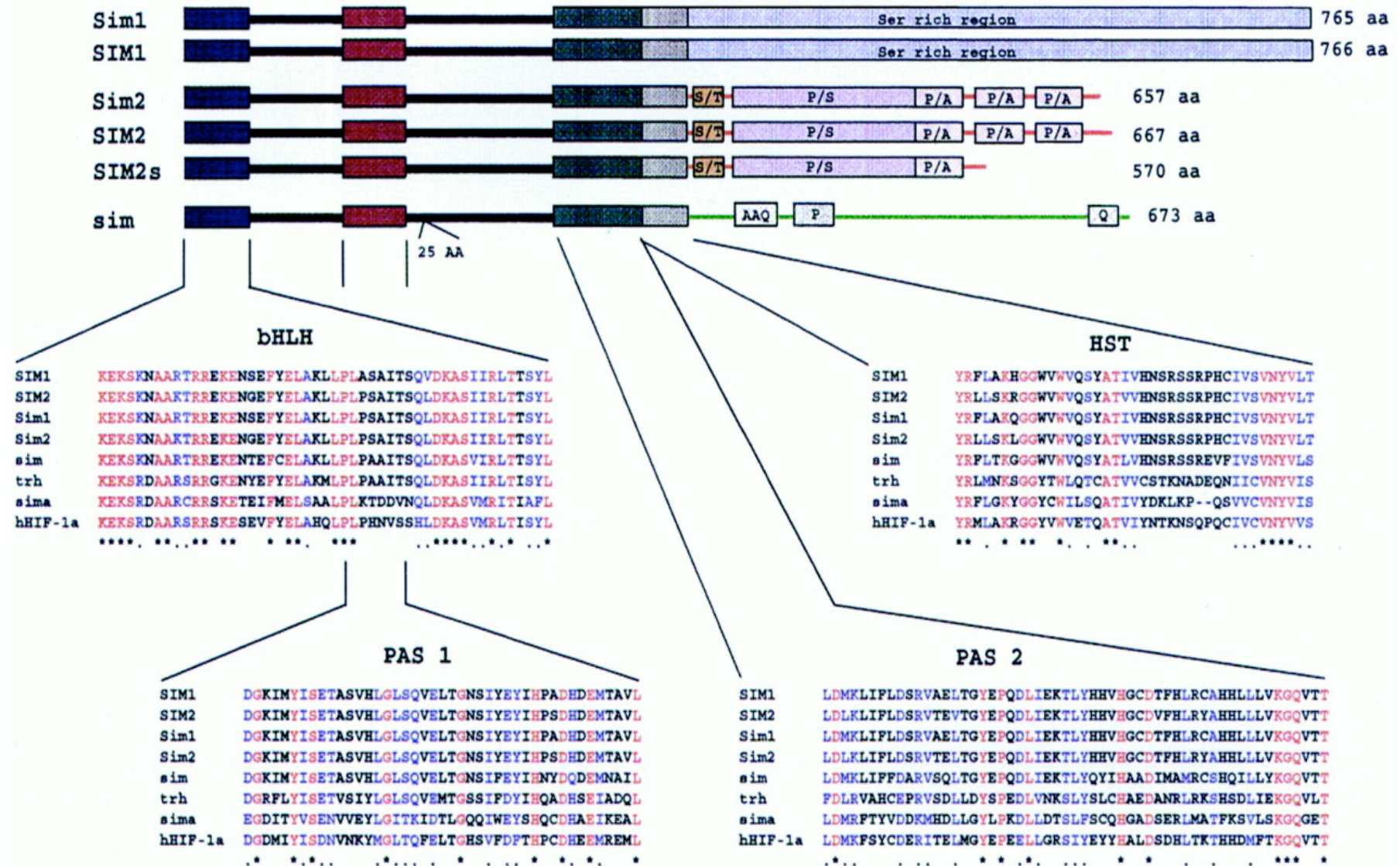
mSim2 expression at E11.5

- ◆ The mapping of hSIM2 on HC21 in the so-called DS critical region
- ◆ Its Drosophila homologue (dsim) is a master regulator of fruitfly neurogenesis expressed early in embryogenesis in the developing midline
- ◆ The expression pattern of its mouse homologue (mSim2) during embryogenesis
- ◆ Transcription factor

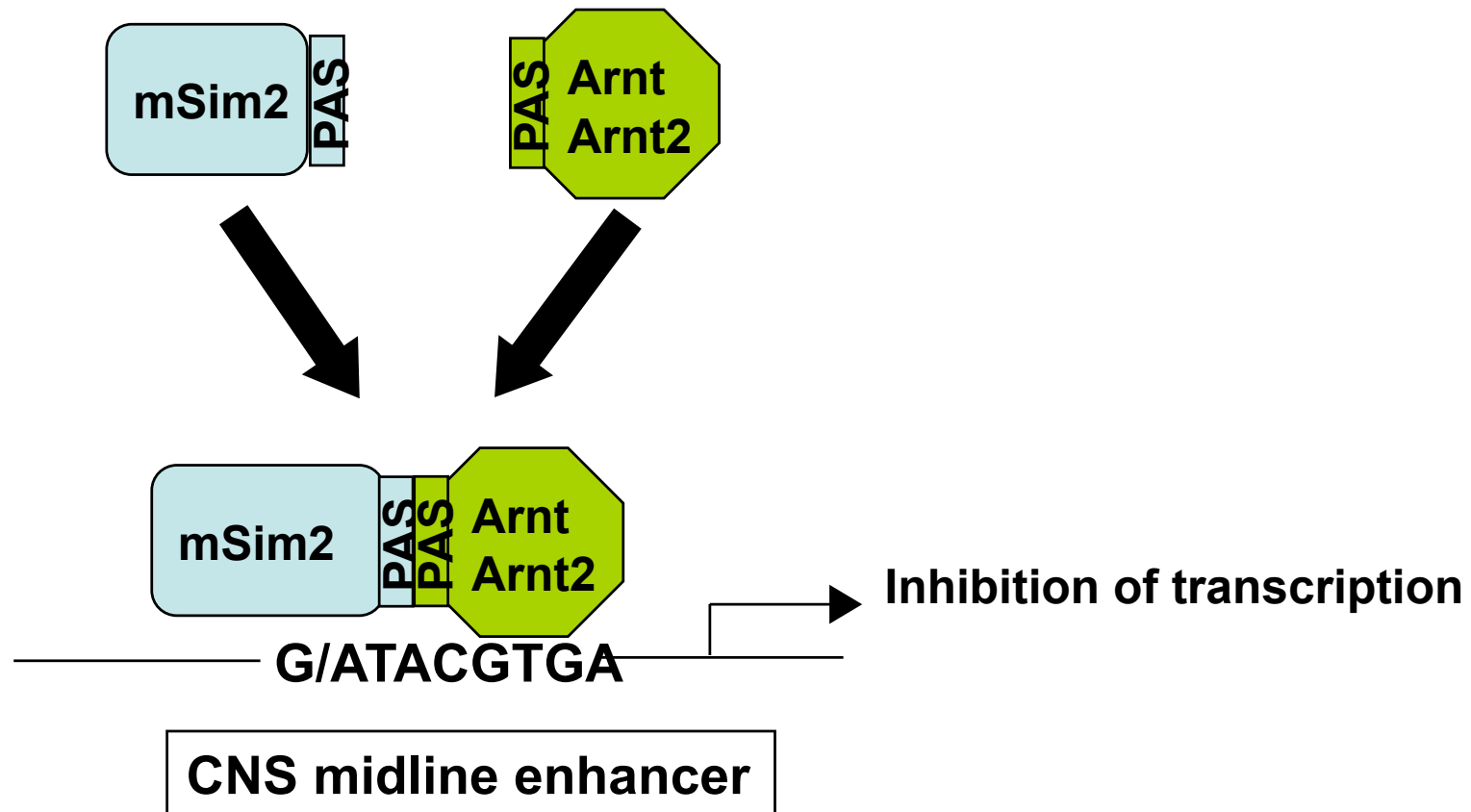


Nambu et al., 1991, Cell 67: 1157-1167.  
Crews et al., 1988, Cell 52: 143-151.

# SIM protein structure

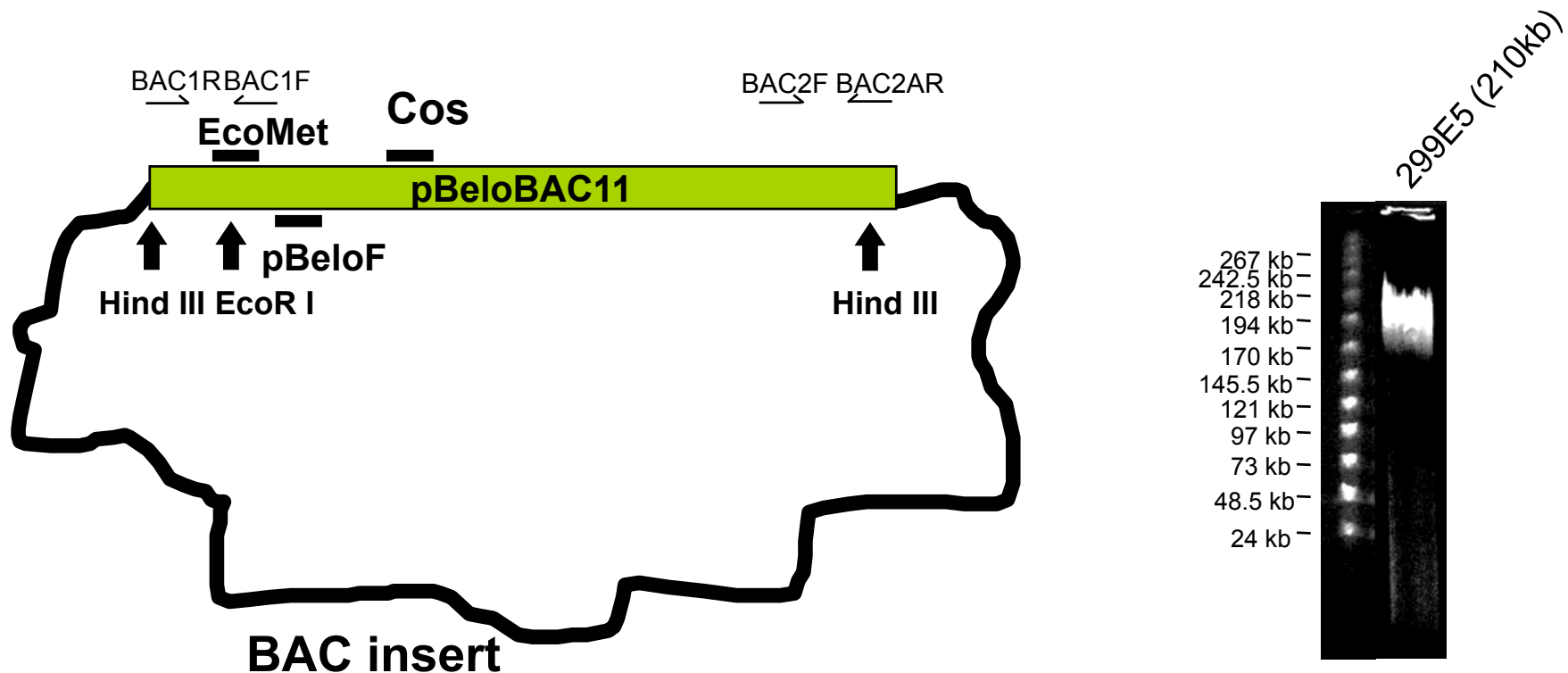


## Sim2 is a transcriptional repressor



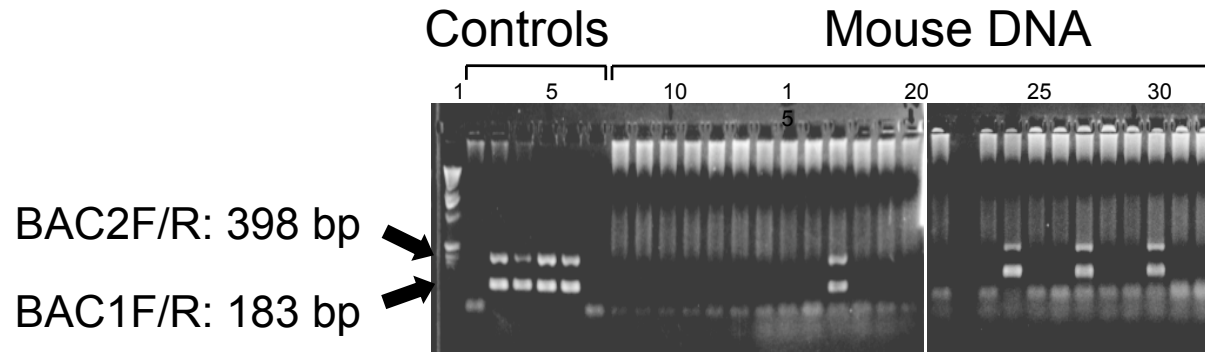
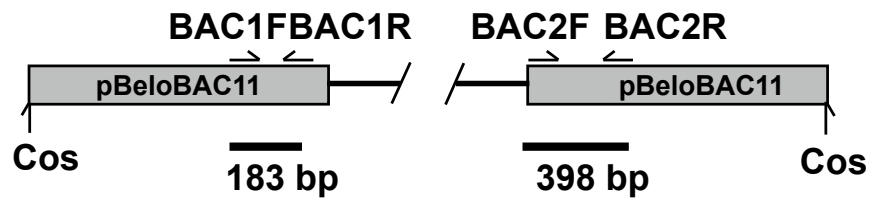
# Isolation and linearization of mSim2 BAC Clone

- ◆ 184N4 (200kb) and 299E5 (210kb) BACs from mouse BAC library
- ◆ Linearization using  $\lambda$  terminase



# Screening of mice for the presence of the mSim2 BAC

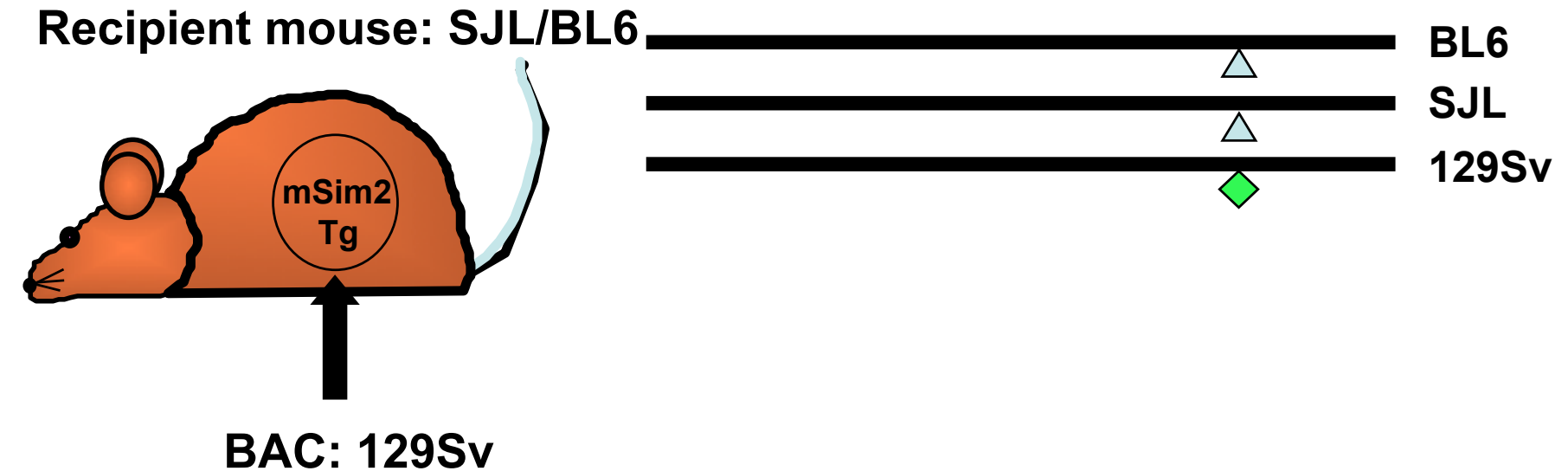
## ◆ Multiplex PCR



**Is the transgene (BAC) derived mSIM2 expressed?**

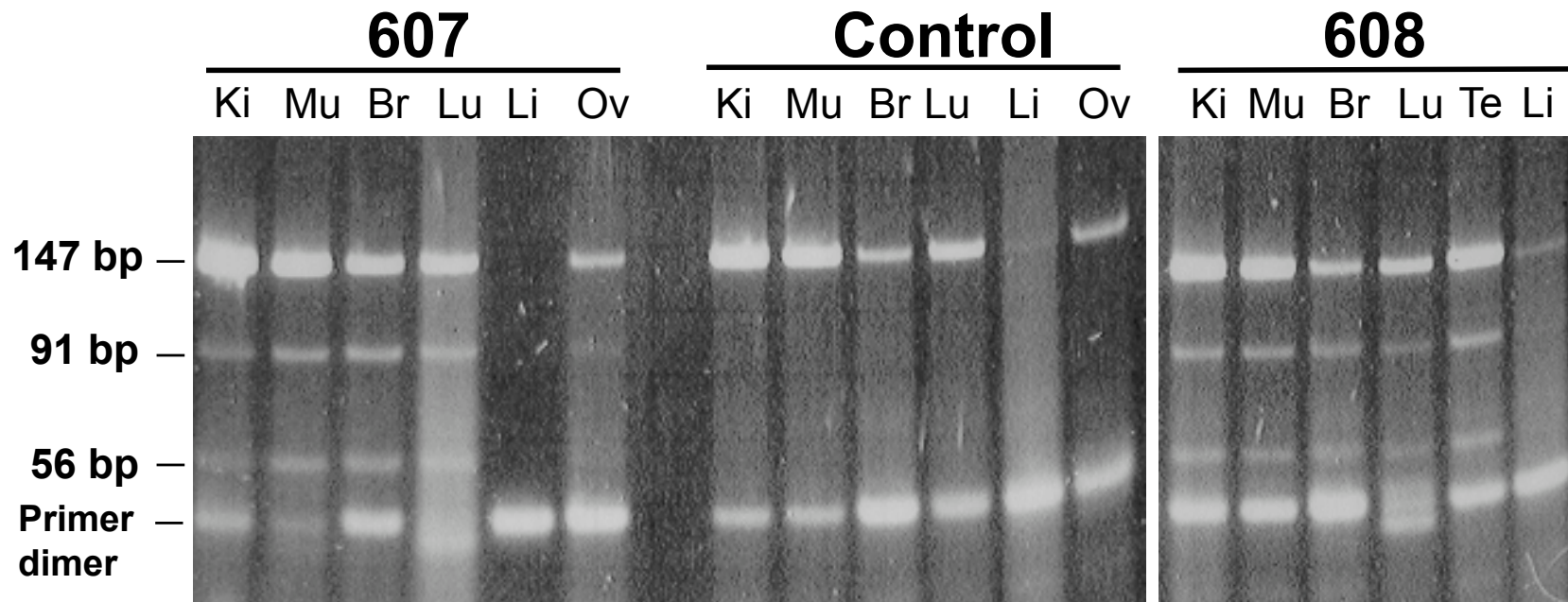


# Identification of a Sim2 polymorphism among inbred mouse strains

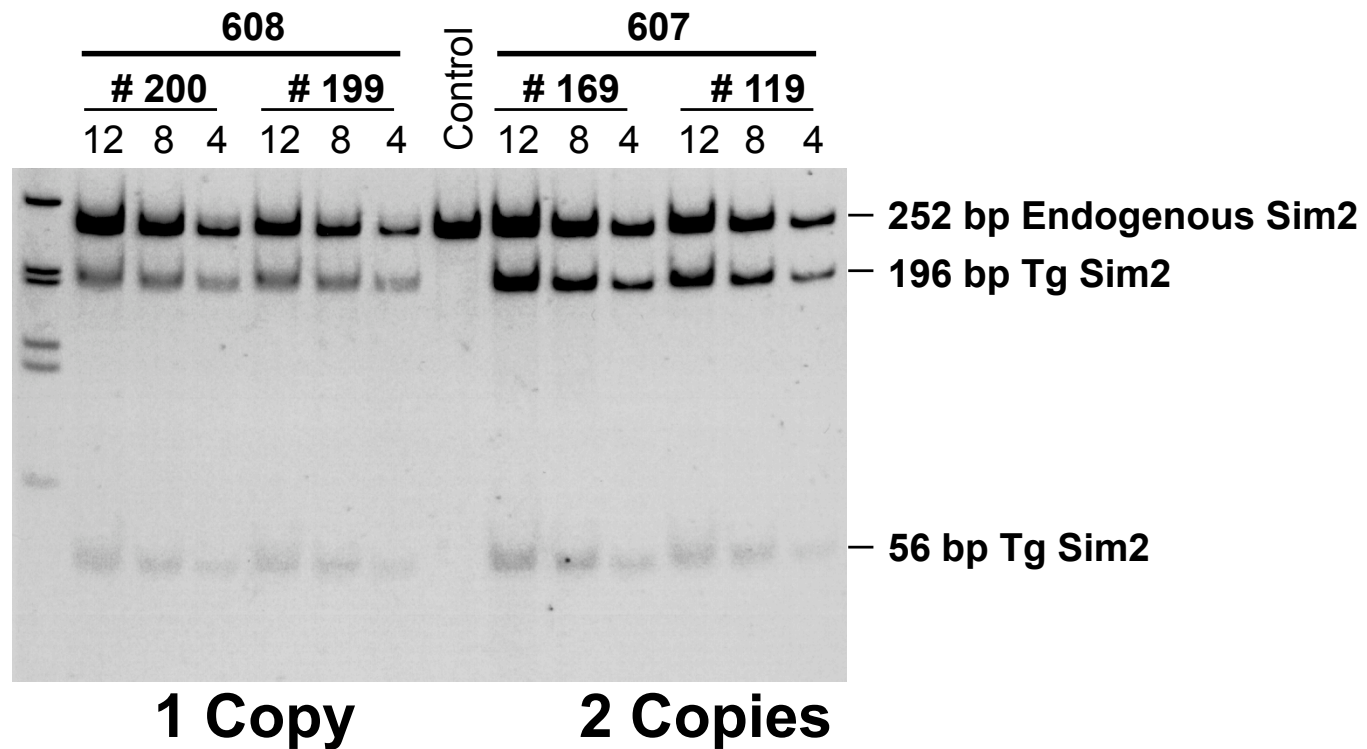
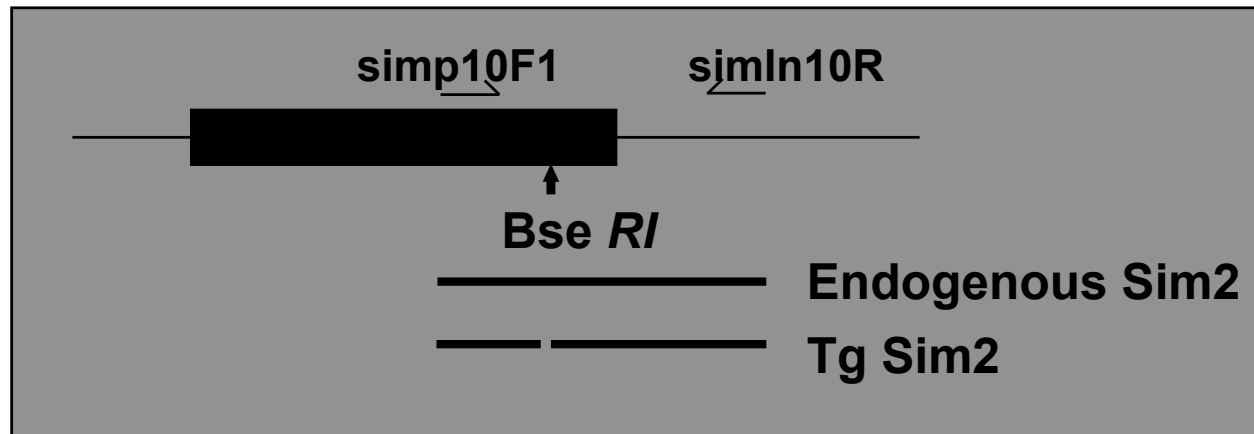


- ◆ Analysis of Tg expression
- ◆ Number of Tg copies

# Detection of the Sim2 BAC transgene expression



# Detection of the Copy Number of the inserted transgene



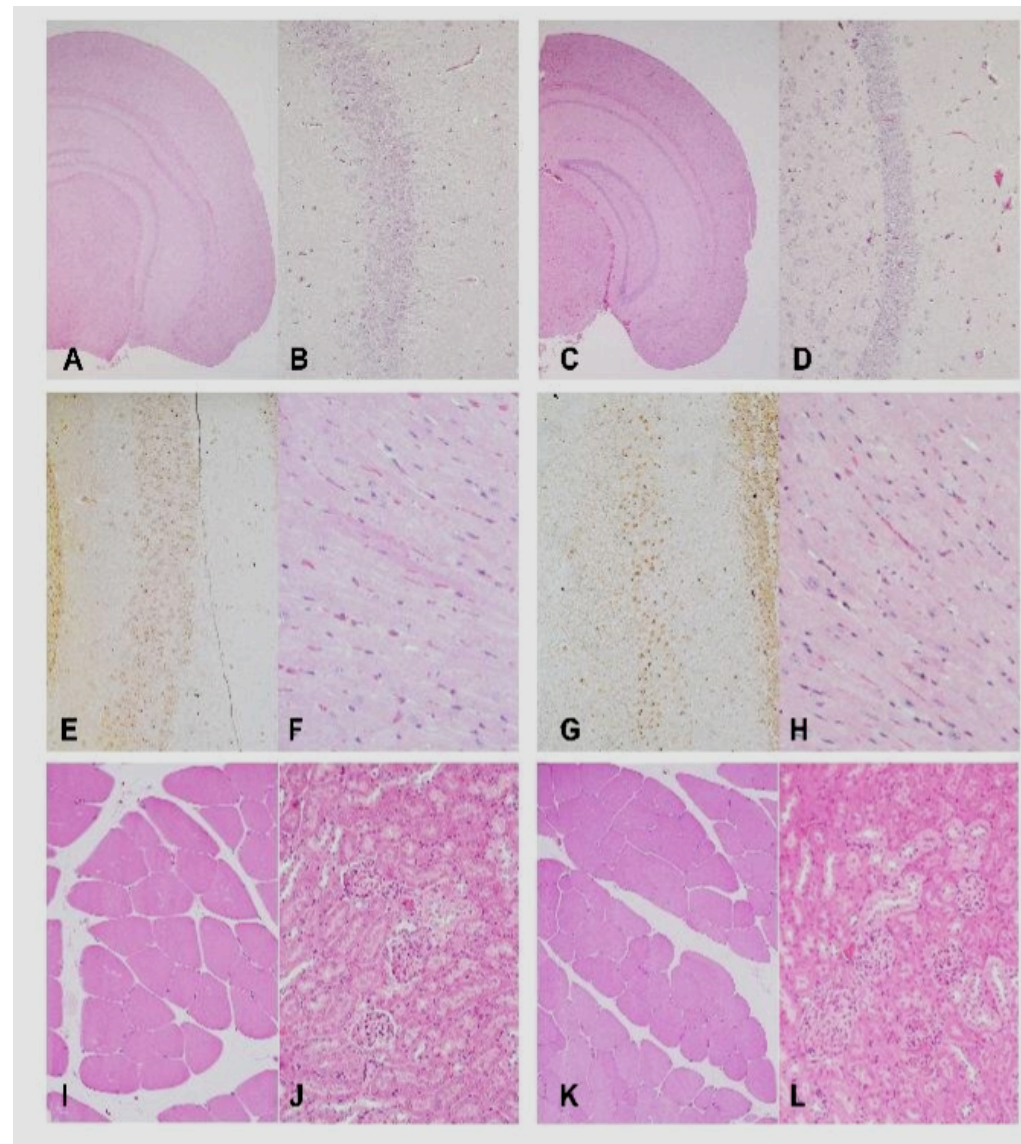
# **mSim2 Tg mice phenotype: general observations**

- ◆ **Transgenic animals from both lines develop normally, are fertile, and do not present any gross detectable phenotypic changes**
- ◆ **Histology**
- ◆ **Behavioural tests:**
  - **Morris water maze**
  - **Fear conditioning**
  - **Open field**
  - **O-maze**
  - **Resident-intruder social interaction test**
  - **Stress-induced analgesia**

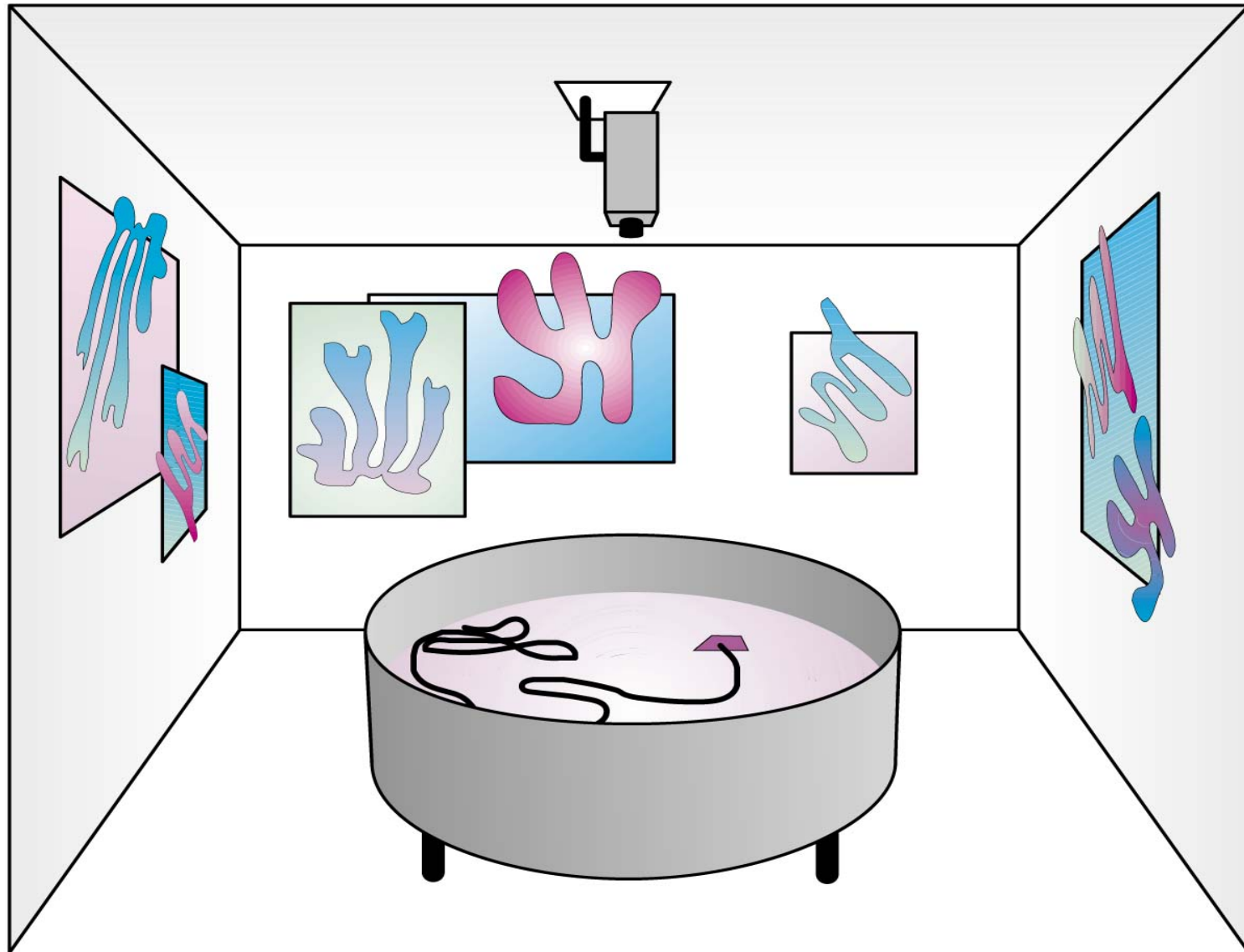
# Histological analysis of Sim2 transgenic mice

**Tg**

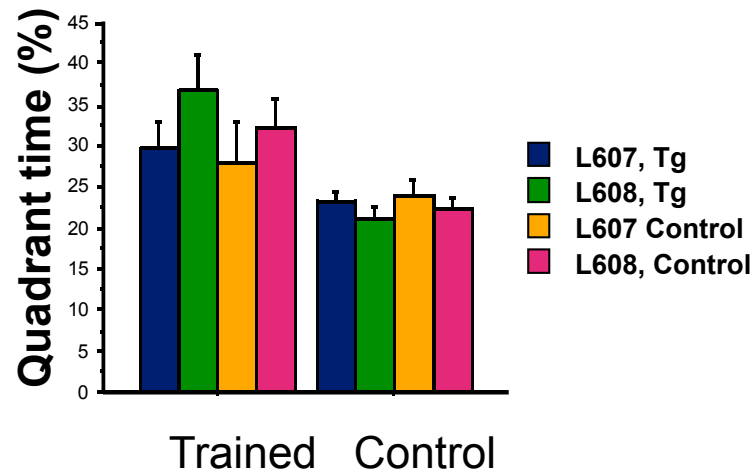
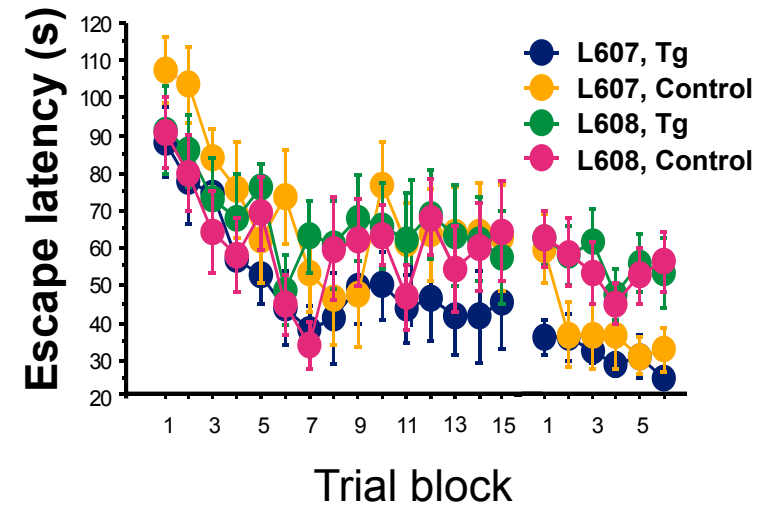
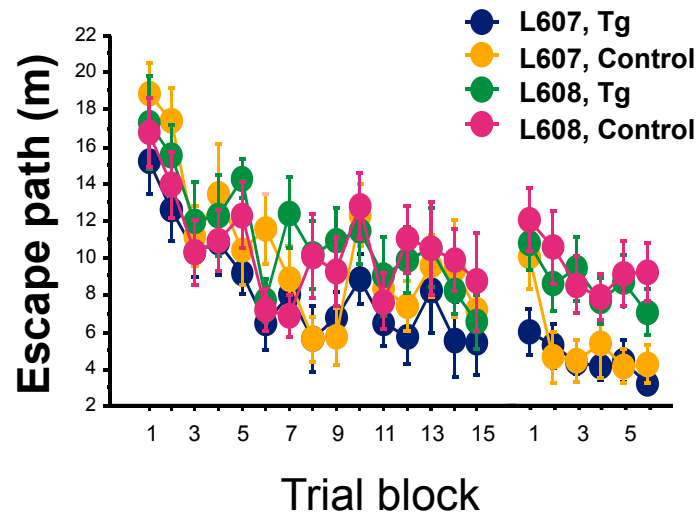
**Control**



# Morris water maze



# Morris water maze test results



Line 607: 12 transgenics and 11 control mice  
Line 608: 14 transgenics and 13 control mice



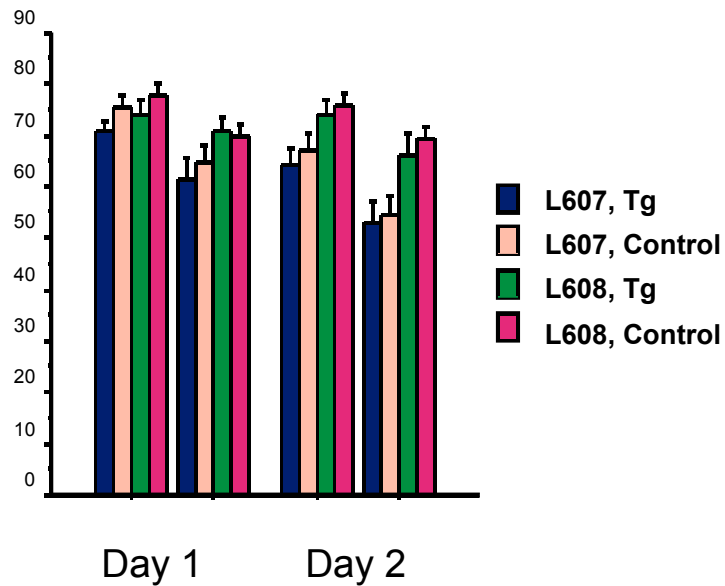
# Open field



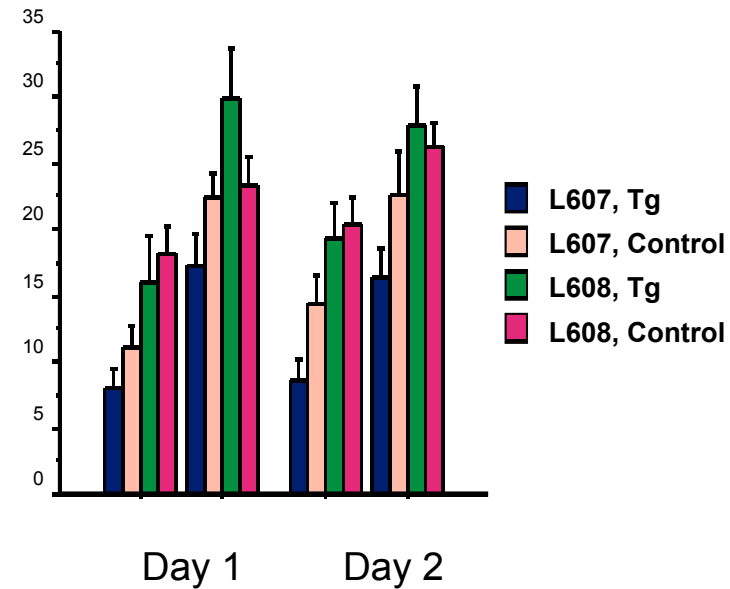


# Open field test results

## Open field: % activity



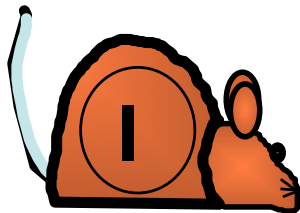
## Open field: % center time



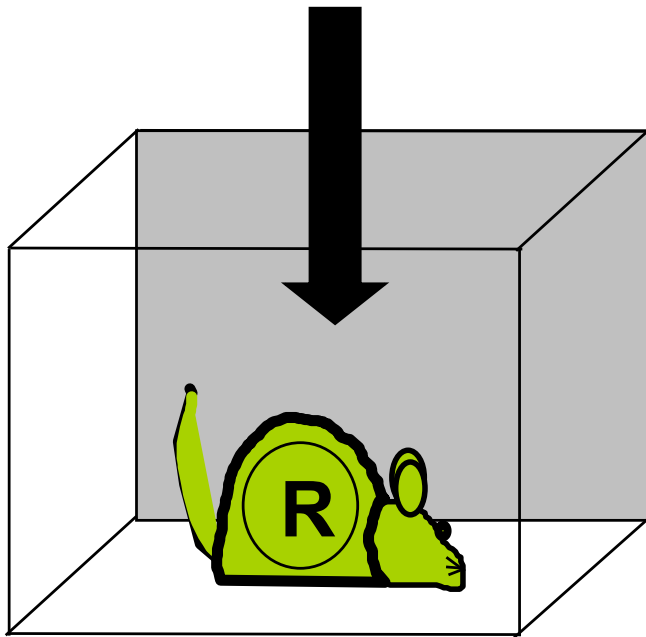
Line 607: 12 transgenics and 11 control mice  
Line 608: 14 transgenics and 13 control mice

# Analysis of social interaction between control and Sim2 transgenic female mice

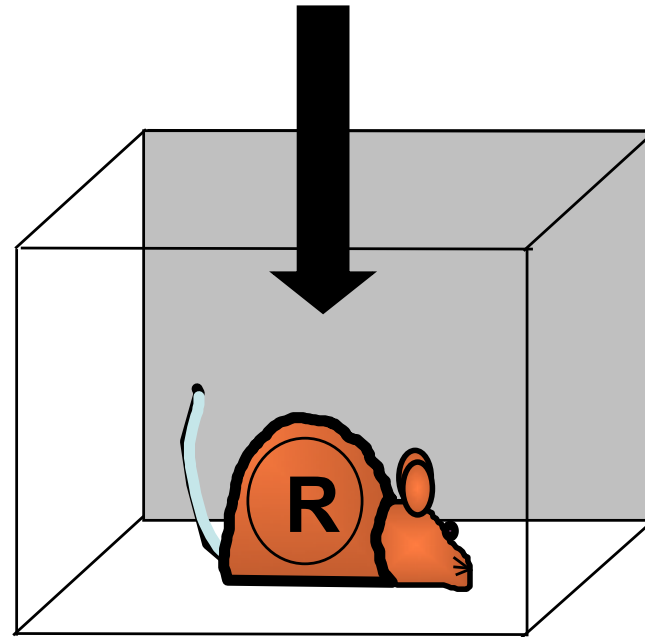
mSim2 Tg



Control

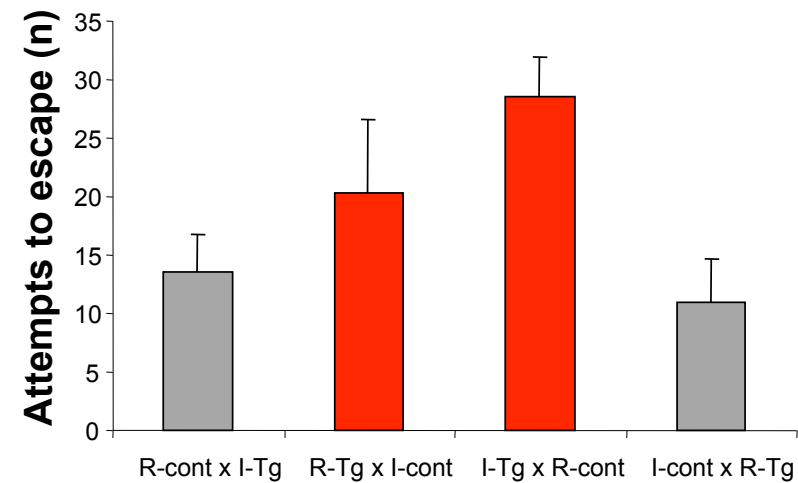
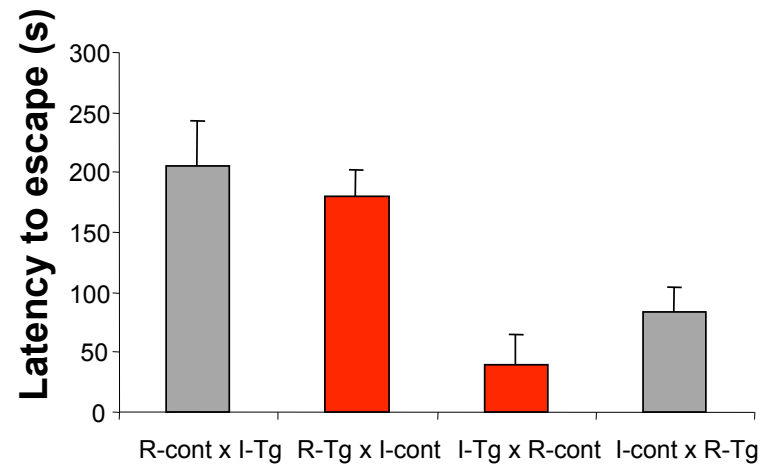


Control

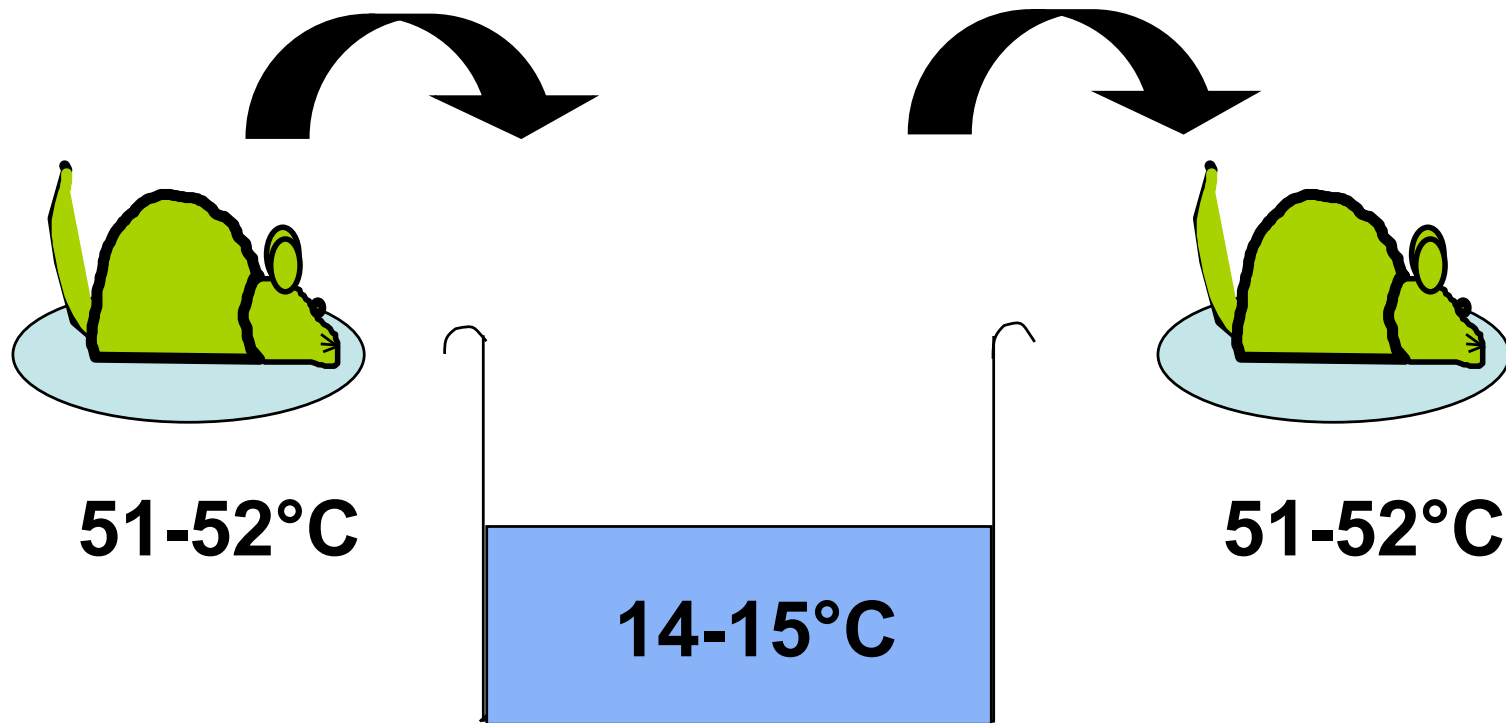


mSim2 Tg

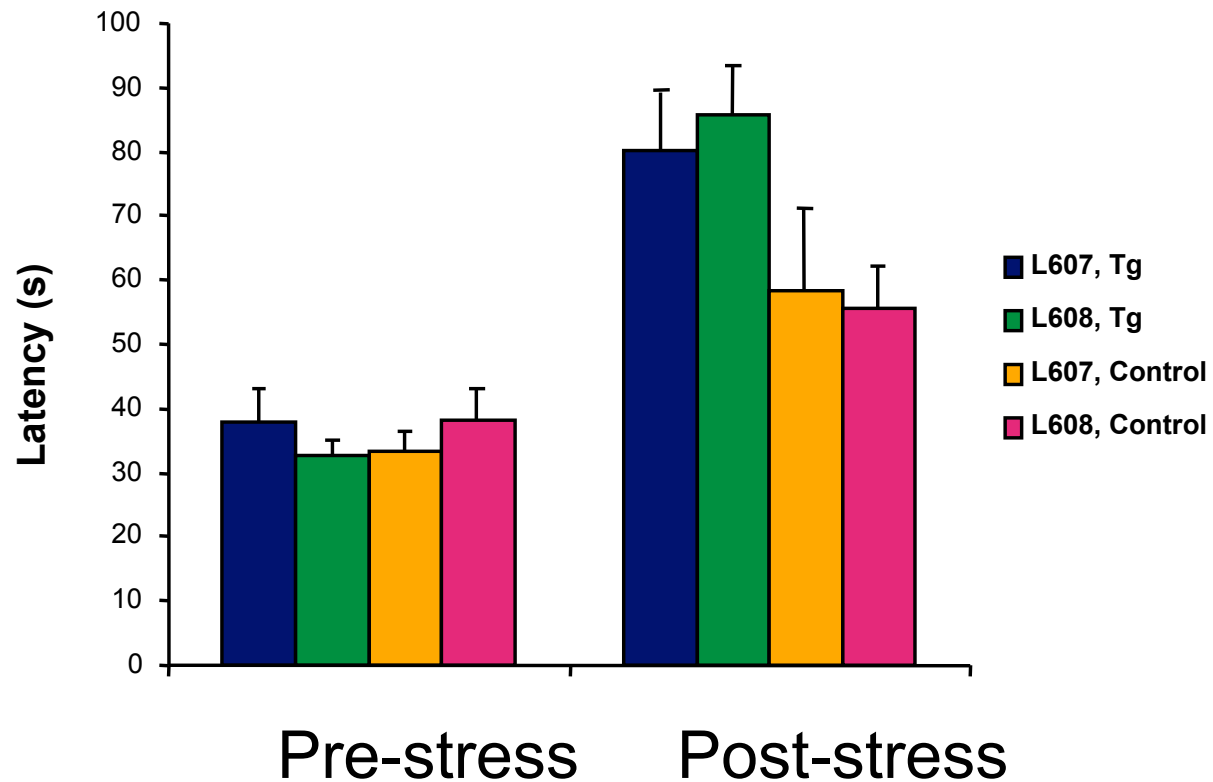
# Analysis of social interaction between control and Sim2 transgenic female mice



# Stress-induced Analgesia Test of Sim2 Transgenic Mice



# Stress-induced Analgesia Test of Sim2 Transgenic Mice



Line 607: 19 transgenics and 10 control mice  
Line 608: 12 transgenics and 18 control mice

# Conclusions

- ◆ Two mSim2 BAC Tg lines created
  - low copy number of the transgene (1-2 copies)
  - correct spatial expression
- ◆ Phenotype
  - Normal growth and fertility
  - No histopathological changes were detected
  - No learning disabilities were detected in MWM and fear conditioning
  - slightly reduced exploratory behaviour
  - Reduced social interaction
  - Reduced responsiveness to pain

Low copy number Sim2 BAC transgenic mice support the hypothesis that the DS phenotypes result from overexpression of a few genes each responsible for certain phenotypes

Human Molecular Genetics, 2000, Vol. 9, No. 12 1853-1864  
© 2000 [Oxford University Press](#)

# **Mice trisomic for a bacterial artificial chromosome with the single-minded 2 gene (*Sim2*) show phenotypes similar to some of those present in the partial trisomy 16 mouse models of Down syndrome**

**Roman Chrast<sup>1,2</sup>, Hamish S. Scott<sup>1</sup>, Rime Madani<sup>3</sup>, Lars Huber<sup>3</sup>, David P. Wolfer<sup>3</sup>, Marco Prinz<sup>4</sup>, Adriano Aguzzi<sup>4</sup>, Hans-Peter Lipp<sup>3</sup> and Stylianos E. Antonarakis<sup>1</sup>**

<sup>1</sup>Division of Medical Genetics, Geneva University Medical School and University Hospital, 1 Rue Michel-Servet, CH-1211 Genève 4, Switzerland, <sup>2</sup>Graduate Program of Cellular and Molecular Biology, Geneva University Medical School, Geneva, Switzerland, <sup>3</sup>Institute of Anatomy, University of Zurich, Zurich, Switzerland and <sup>4</sup>Institute of Neuropathology, University Hospital of Zurich, Zurich, Switzerland

*Proc. Natl. Acad. Sci. USA*  
Vol. 95, pp. 6256–6261, May 1998  
Genetics

## **Ts1Cje, a partial trisomy 16 mouse model for Down syndrome, exhibits learning and behavioral abnormalities**

HARUHIKO SAGO\*, ELAINE J. CARLSON\*, DESMOND J. SMITH<sup>†</sup>, JOSHUA KILBRIDGE<sup>‡</sup>, EDWARD M. RUBIN<sup>†</sup>,  
WILLIAM C. MOBLEY<sup>‡</sup>, CHARLES J. EPSTEIN\*, AND TING-TING HUANG\*§

Departments of \*Pediatrics and <sup>‡</sup>Neurology, University of California, Box 0546, San Francisco, CA 94143-0546; and <sup>†</sup>Human Genome Center, Lawrence Berkeley National Laboratory, Berkeley, CA 94720



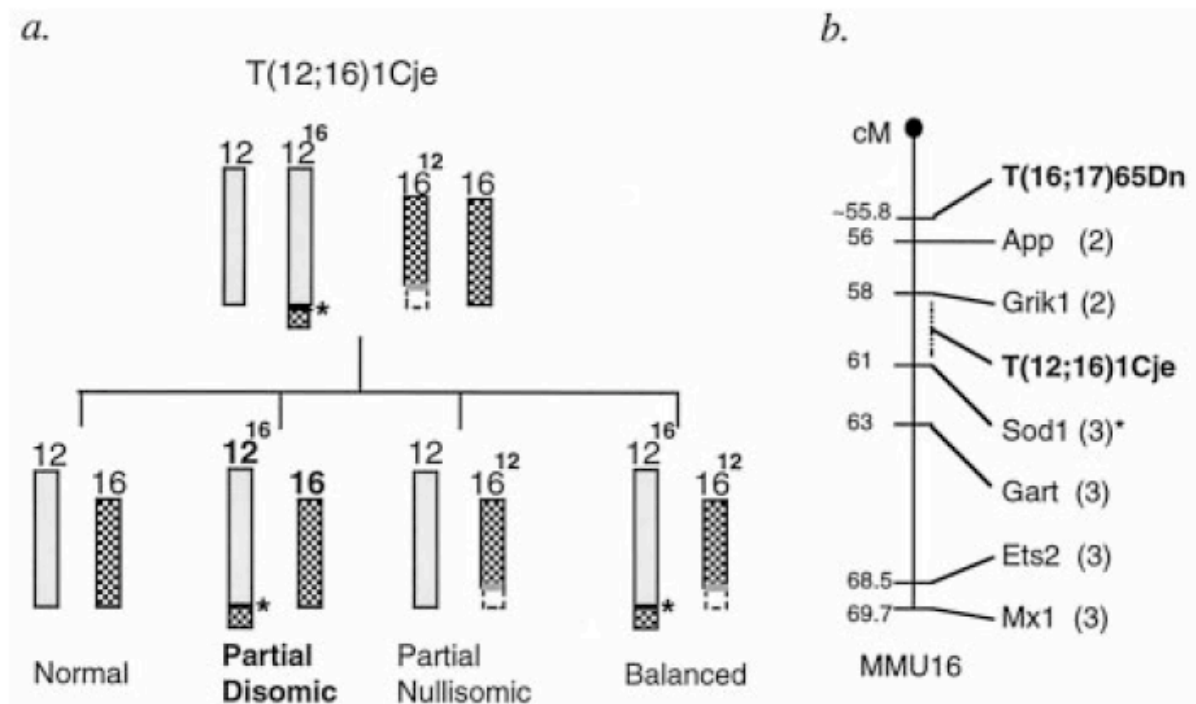
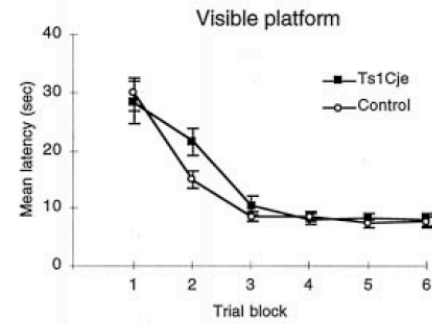
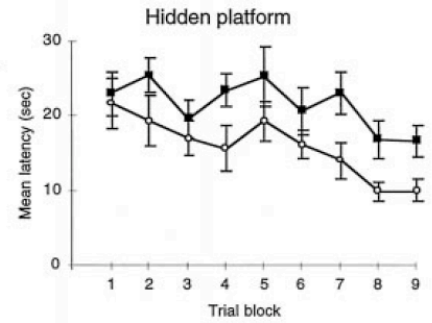


FIG. 2. (a) Diagram of meiosis of the 12;16 reciprocal translocation in the heterozygous *Sod1* mutants. Four types of gametes are obtained: normal, disomic, and nullisomic for the distal region of MMU16, and balanced. The partial disomic gametes produce a partial trisomy 16 (Ts1Cje). \* indicates the neomycin resistance marker on the 12<sup>16</sup> translocation. (b) Diagram of the mapping of the triplicated region in Ts1Cje by FISH analysis. The numbers in the parentheses indicate the copy numbers of genes detected by FISH. There is segmental trisomy of the region from *Sod1* to *Mx1*. \*, *Sod1* is not functionally trisomic because the *Sod1* gene in the translocated segment is inactivated by the insertion of the neomycin resistance sequence.

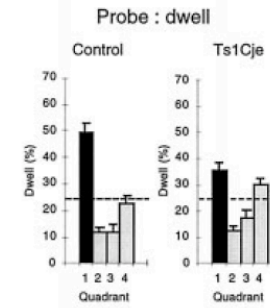
*a.*



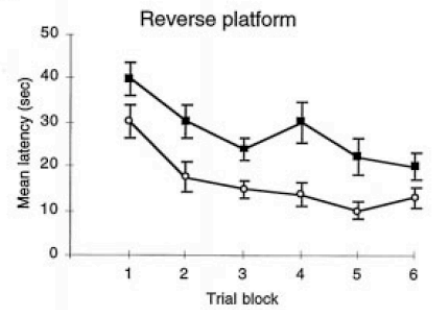
*b.*



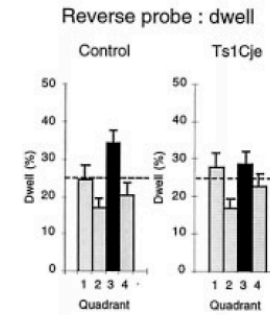
*c.*



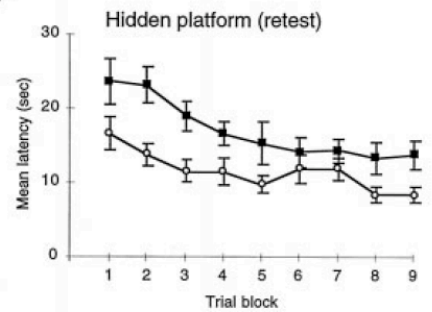
*d.*



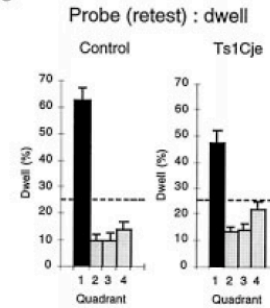
*e.*



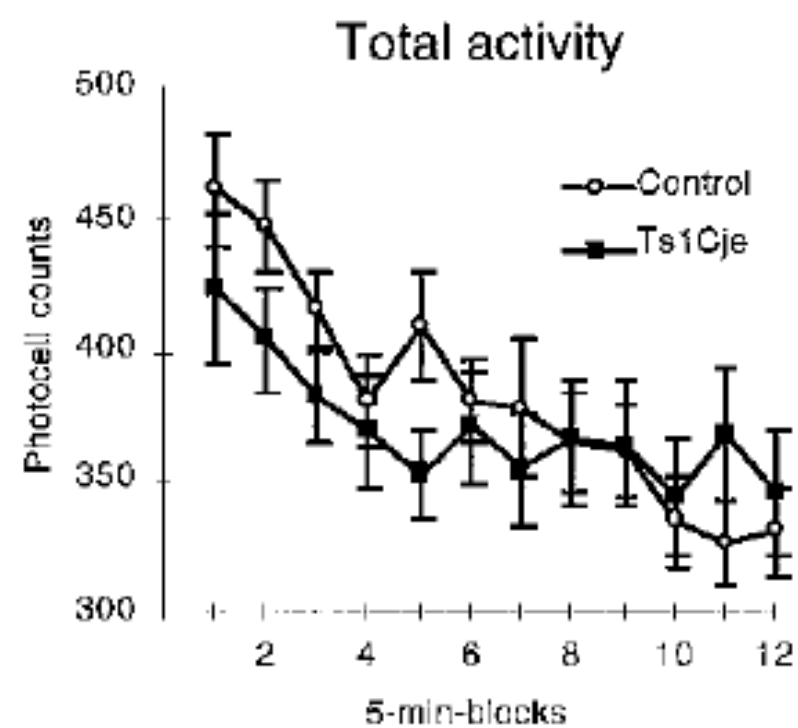
*f.*



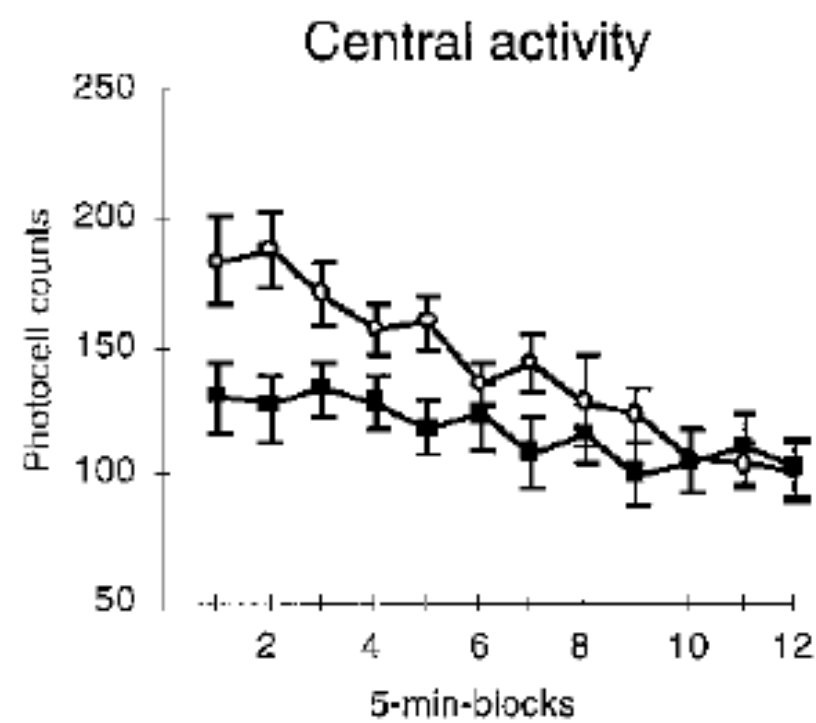
*g.*



*a.*



*b.*

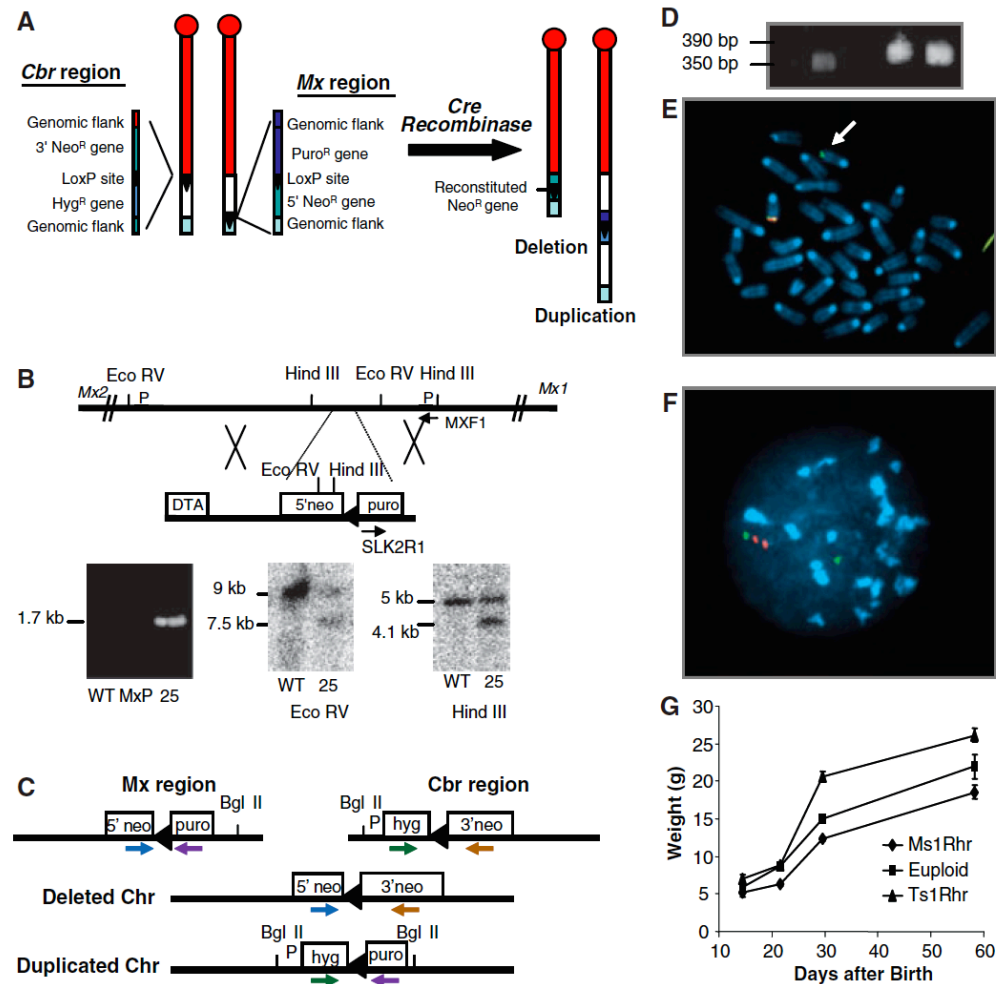


	Ts65Dn	Ts1Cje	YAC152F7tel	YAC230E8
<b>Triplcation</b>				
<i>App</i>	●			
<i>Sod1</i>		○		
<i>Mnbh</i>			●	●
<i>Mx1</i>	●	●		
MMU16	15 cM	10 cM	180 kb	670 kb
	+ cent MMU17		(human)	(human)
<b>Learning deficits</b>				
Visible (non-spatial)	+	-	-	-
Hidden (spatial)				
Platform	+	+	-	-
Probe	+	+	+/-	-
Reverse ("flexibility")	ND	+	+	+
BFCN loss (6 months)	+	-	ND	ND

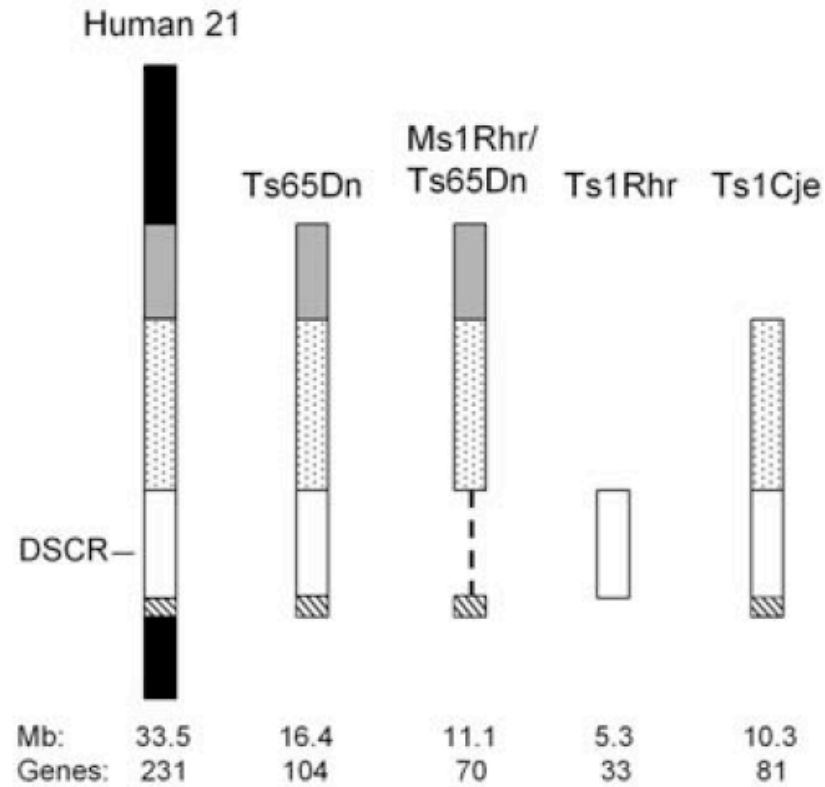
# A Chromosome 21 Critical Region Does Not Cause Specific Down Syndrome Phenotypes

L. E. Olson,<sup>1\*</sup> J. T. Richtsmeier,<sup>2</sup> J. Leszl,<sup>2</sup> R. H. Reeves<sup>1†</sup>

The "Down syndrome critical region" (DSCR) is a chromosome 21 segment purported to contain genes responsible for many features of Down syndrome (DS), including craniofacial dysmorphology. We used chromosome engineering to create mice that were trisomic or monosomic for only the mouse chromosome segment orthologous to the DSCR and assessed dysmorphologies of the craniofacial skeleton that show direct parallels with DS in mice with a larger segmental trisomy. The DSCR genes were not sufficient and were largely not necessary to produce the facial phenotype. These results refute specific predictions of the prevailing hypothesis of gene action in DS.

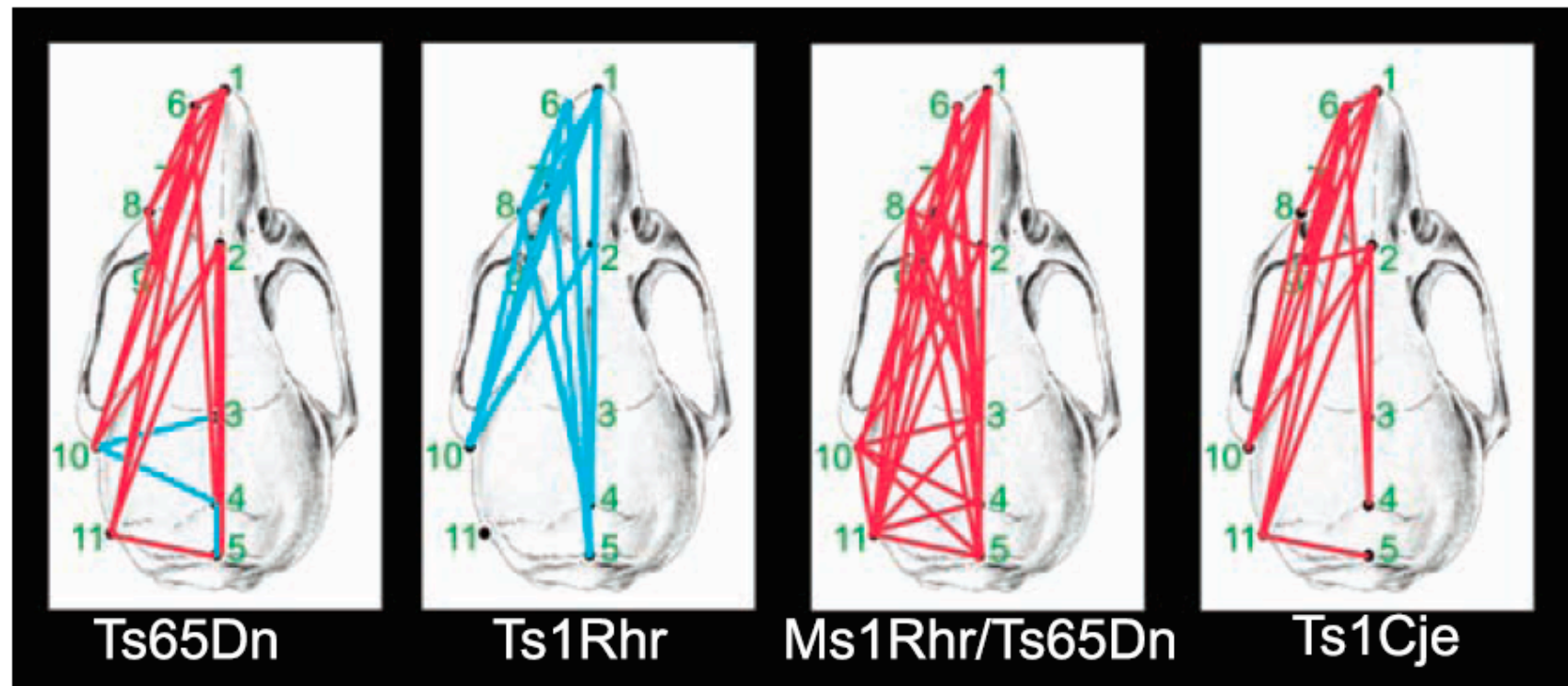


**Fig. 1.** Construction of a duplication or deletion of the MMU16 region orthologous to the DSCR. (A) LoxP sites were targeted to asymmetrical positions on MMU16 at *Cbr1* (13) and adjacent to *Mx2*. Each targeting vector contained a LoxP site (triangle), a selectable antibiotic resistance gene (hyg or puro), and half of the neomycin resistance gene (5' or 3' neo). (B) The *Mx*-Lox vector produced a 1.7-kb PCR product from the targeted allele in line MxP25. Wild-type (WT) and targeted alleles produced 9.0- and 7.5-kb restriction fragments with *EcoRV* and 5.0- and 4.1-kb fragments with *HindIII*. Arrows identify PCR primers; P designates probes. (C) PCR primers used to screen vector sequences for recombination after Cre-mediated translocation. (D) PCR products from *neo<sup>r</sup>* ES lines produced the 350- and 390-base pair (bp) fragments expected for deletion and duplication, respectively. (E) Metaphase FISH with one bacterial artificial chromosome (BAC) that maps to the DSCR (red) and a second BAC proximal to it (green) shows one chromosome with a single green signal (arrow) and a second with green and red plus a yellow signal, indicating overlap. (F) Interphase FISH shows a green signal by itself representing the deleted MMU16 and a green signal with two adjacent red signals representing the duplication. (G) Body weights of Ts1Rhr mice are significantly larger than controls. Standard errors (bars) are indicated.



**Fig. 2.** Trisomic segments represented in mouse models relative to HSA21. The DSCR is indicated as an open box where it is present and by a dashed line where it is deleted. Approximate size in megabases of triplicated information is based on genomic sequence; gene number is from Gardiner *et al.* (11). Gene content for each segment is given in table S1.





**Fig. 3.** Linear measurements that contribute to dysmorphology in Ts65Dn, Ts1Cje, and Ms1Rhr/Ts65Dn form a closely related set, whereas trisomy for the DSCR alone produces a distinct dysmorphology in Ts1Rhr. Red lines indicate linear distances between landmarks that were significantly smaller in trisomic mice relative to euploid mice; blue lines indicate distances that were larger. Statistical significance was determined using Euclidean Distance Matrix Analysis confidence intervals (17). This is not a simple scaling difference because the magnitude of the differences varies from one measure to another and the set of differences contributing to dysmorphology in Ts1Rhr is different from that in the other models. Landmarks are defined in fig. S3.